DESIGN AND SYNTHESIS OF SMALL MOLECULES TO TARGET PROTEIN-PROTEIN INTERACTIONS

A Dissertation

by

DONGYUE XIN

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Kevin Burgess
Committee Members,	Daniel A. Singleton
	Tadhg Begley
	Thomas R. Ioerger
Head of Department,	François Gabbaï

May 2016

Major Subject: Chemistry

ABSTRACT

Protein-protein interactions (PPIs) play key regulatory roles in biological systems, and some of these are interesting drug targets. Consequently, it is important to develop generally applicable methods to identify small molecules that disrupt or disturb PPIs. One emerging approach is to use novel small molecule scaffolds to mimic protein-protein interfaces. To identify good mimics for PPI targets, a novel computational approach *Exploring Key Orientations (EKO)* has been developed. This thesis is focused on the design, synthesis, EKO analysis and biological applications of two interface mimic scaffolds.

The first mimic, an oligo-piperidine-piperidinone (OPP) scaffold was designed and synthesized to target extended interface regions. Derivatives of this scaffold have been efficiently prepared in a divergent-convergent method. Conformational studies revealed that the OPP scaffold could exist in an extended helical conformation and it was a good mimic of ideal α-helix in solid state. Further investigations by molecular modeling indicated this scaffold could be a multi-faceted mimic for several secondary structure motifs in solution. An interesting protein target called antithrombin was discovered with EKO database mining analysis for the OPP scaffold. In biological studies, derivatives of OPP were found to interfere with oligomerization of antithrombin in a side-chain and concentration dependent manner.

As an orthogonal interface mimic, a new constrained cyclic peptide-organic hybrid was also explored to target compact PPI interface regions. An anthranilic acid was incorporated in the scaffold as a turn-inducing motif. Extensive conformational analyses by 1D and 2D NMR, CD, and molecular modeling were performed and the results showed that the cyclic peptide scaffold could mimic multiple turn structures. Moreover, these new turn mimics were conformationally homogeneous in solution and their conformations had a strong and predictable correlation with side-chain stereochemistries.

The scaffolds described in this thesis represent suitable scaffolds to target protein-protein interactions. Compared with traditional methods, interface mimicry approach together with EKO analysis can significantly facilitate the discovery of small molecules for protein-protein interactions.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Kevin Burgess for his guidance and support during the course of this research. I have benefited a lot from his advices and his passion for science. Thanks also go to my committee members, Dr. Daniel A Singleton, Dr. Tadhg Begley and Dr. Thomas R. Ioerger for their insightful comments and suggestions throughout my PhD study.

I am also grateful for working with so many talented coworkers and collaborators over the years. Special thanks to Dr. Lisa M. Pérez and Dr. Thomas R. Ioerger for their guidance and helpful discussions on molecular modeling; to Dr. Andreas Holzenburg for his assistance on TEM; to Dr. Joseph H. Reibenspies for his help with X-ray crystallography; to Dr. Arjun Raghuraman, Dr. Dmytro Fedoseyenko and Dr. Eunhwa Ko for their mentorship and help when I first joined Dr. Burgess's research group; and to all the Burgess group members, all my friends and the department faculty and staff for making my life at Texas A&M University a wonderful experience.

Finally, thanks to my parents and my parents-in-law for their love, encouragement and understanding. I feel so lucky to meet Xiao Cong in college eight years ago. She is my wife, my best friend, my favorite contact on my cellphone and my personal consultant who I can fully trust with my life. I am thankful for sharing the entire amazing journey with her and her love is my best source of strength and happiness.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

LIST OF SCHEMES

Page

CHAPTER I

INTRODUCTION*

1.1 Protein-protein Interactions as Drug Targets

Protein-protein interactions (PPIs) are widely involved in cell-signaling, gene expression regulation, immune response initiation and many processes of cell metabolism.[1,2] Since many of these interactions play critical regulatory roles in biological processes such as cell proliferation and apoptosis, it is of great significance to study these interactions to have a better understanding of the biological systems. Specifically, a lot of these interactions are related to many diseases such as cancer, Alzheimer's disease, diabetes and high cholesterol.[3-5] Therefore, studies towards selective modulation of the function of protein-protein interactions at the molecular level may lead to breakthrough therapeutics for some of the most challenging diseases.[6-8] Some of the most widely studied PPIs are highly pharmaceutically interesting such as p53/MDM2[9] and Bak/Bcl2 interactions,[10] which are related to cell apoptosis and thus potential cancer targets.[1,11,12] Consequently, it is highly desirable to develop generally applicable methods to identify molecular entities that are able to disrupt PPIs.

---

*Reprinted in part with permission from "Extended Piperidine-piperidinone Protein Interface Mimics", Dongyue Xin, Arjun Raghuraman and Kevin Burgess, *J. Org. Chem.* **2015**, *80*, 4450-4458. Copyright 2015 American Chemical Society and from "Small Molecule Probes That Perturb a Protein-protein Interface in Antithrombin", Dongyue Xin, Andreas Holzenburg and Kevin Burgess, *Chem. Sci.* **2014**, *5*, 4914-4921. Copyright 2014 The Royal Society of Chemistry.

1.2 Current Methods to Target Protein-protein Interactions

Protein-protein interactions can be targeted by monoclonal antibodies or small molecules. Significant theoretical and experimental techniques have been developed in the past two decades to increase the success rate and reduce the cost to identify novel molecules for selected PPIs. However, compared with enzymatic targets, it is still much more challenging to develop therapeutics for PPI targets.[1]

Monoclonal antibodies (mAbs) are well suited in targeting certain types of PPIs, especially the ligand-cell surface receptor interactions such as EGF/EGFR[13,14] and PD1/PDL1[15,16] and this led to some revolutionary treatments for certain types of cancer. However, antibodies also have some limitations in biomedicine. MAbs are macromolecular proteins, so most are generally not cell permeable and thus their current applications are mainly limited to cell surface receptors. Furthermore, mAbs do not tend to penetrate into solid tumors[17,18,19,20] because they are too big to leave blood vessels and efficiently diffuse into tissue. MAbs that *do* diffuse into tissue tend to be trapped by antigens located on the perivascular tumor cells, preventing permeation into the tumor mass,[21] *ie* there is an "antigen barrier",[22-24] even around micrometastases.[22] Moreover, slow clearance of mAbs from the body results in high normal tissue exposure[25-27] hence most mAb molecules in a dose accumulate in the excretory organs (intact mAbs in the liver, fragments in the kidneys) and do *not* reach their targets.[28,29] MAbs also can be immunogenic, even when they are humanized or human;[30] this can cause hypersensitivity, neutralizing effects, and changeable pharmacokinetic properties.[29] These limitations should be considered alongside problems with non-specific

conjugation chemistry leading to reduced product homogeneity,[31] cost, and stability/shelf life issues. Despite these factors, interest in mAb has surged *because it is comparatively easy to raise antibodies to cell surface receptors*.[32]

The most frequently used method to find small molecules for PPI targets is high throughput screening (HTS), which is widely applied in the development of inhibitors for enzymes.[33] However, HTS for PPI targets usually gives disappointing hit-rates. It has been suggested that one reason is that the compound collections screened do not have appropriate chemotypes.[1] Besides HTS, there are only a few methods suitable to develop small molecules for PPI targets. Fragment-based approaches using NMR or X-ray supported by modeling, and "tethering", are exciting but the process of joining small pharmacophores into a small molecule that retains the correct shape is difficult and not always attainable.[34,35,36] Some other methods such as the anchor approach[37,38] provides only partial solution for specific PPIs. In addition, in many cases, the lack of small and well-characterized binding pockets makes it difficult to use normal docking programs for virtual screening.[1]

One generally accepted approach specifically targeting PPIs is using secondary structure mimics.[8,39,40] Peptides corresponding to the interface regions are usually poor starting points because PPIs usually involve discontinuous amino-acid residues and peptides are vulnerable to proteases and have other undesirable pharmacological properties.[1,40] To overcome these problems, several semi-rigid molecular scaffolds that can present side chains corresponding to interface residues have been developed. Pioneering work by Hamilton[41-43] and others[42,44,45] featured biphenyl, terphenyl and

related scaffolds to form what we call *minimalist peptidomimetics.*[46,47] These compounds do not have polyamide backbones, are impervious to proteases, and can be cell permeable.

1.3 Minimalist Mimics

Minimalist mimics of peptides and proteins are organic scaffolds that: (i) access fewer conformations than peptides (*ie* are "semi-rigid"); and, (ii) display three or more amino acid side-chains. Motivation for studies of minimalist mimics stem from their potential to perturb protein-protein interactions (PPIs). Non-covalent interaction energies at protein interfaces tend to be dominated by side-chain to side-chain interactions.[48,49] A minimalist mimic that presents side-chains in orientations corresponding to one PPI component, called here the *protein-ligand*, has the potential to displace that ligand and bind to the other partner in the PPI, the *protein-receptor.* (Figure 1.1)

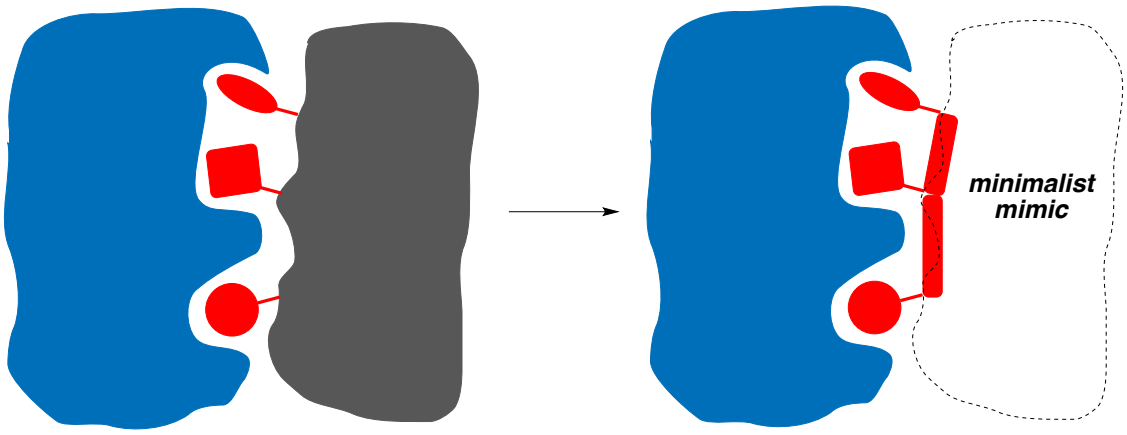


Figure 1.1. Mimicking side-chain orientations at PPI interfaces with minimalist mimics.

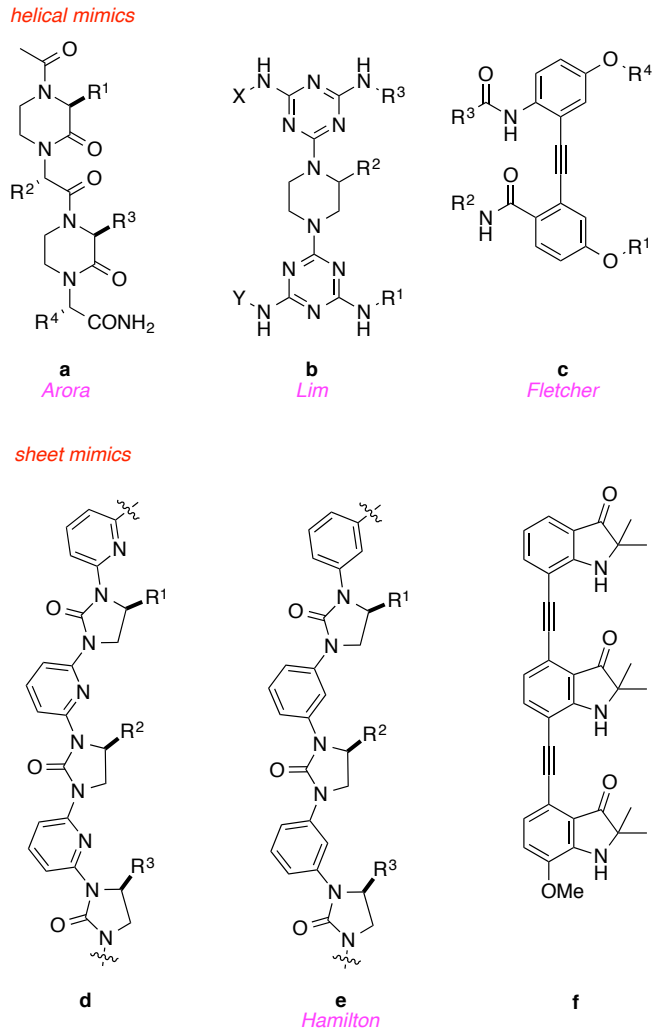


Figure 1.2. Recent examples of minimalist mimics.

Figure 1.2 shows illustrative minimalist mimics (**a**,[45,50,51] **b**,[52] **c**,[53] **d**,[54] **e**,[55] **f**[56]), reported since this area was reviewed in 2011.[47] For any particular PPI, a minimalist mimic scaffold only has potential if it is synthetically accessible with amino acid side-chains corresponding to the protein-ligand at the interface region. The main shortcoming of designing generic secondary structure mimics, is that secondary

structures at PPI interfaces are usually distorted, and many interfaces do not involve secondary structures at all. This makes it difficult to decide where to overlay the mimic on the interface region, and therefore what amino acids side-chains should be incorporated.

If the appropriate region of overlay can be determined, the scaffold has to be made presenting those particular side-chains. Mimic **a**, has been formed from amino acids, and is attractive in this regard. Sheet mimics **d** – **f** have only so-far been prepared with methyl side-chains, but there is potential to construct **d** and **e** from a variety of amino alcohols. Preparation of mimics **b**, **c**, and **f** with a variety of genetically encoded amino acid side-chains would be more challenging.

In our view, the hidden potential of so-called minimalist mimics is they *also* populate conformations other than the targeted secondary structure.[46,47,57] (Figure 1.3) This led us to appreciate that minimalist mimics in some situations can resemble regions of a PPI interface but *not* any particular secondary structure. This was enlightening because it forced us to realize that mimicry of secondary structures is not the real issue in designing small molecules to perturb PPIs.

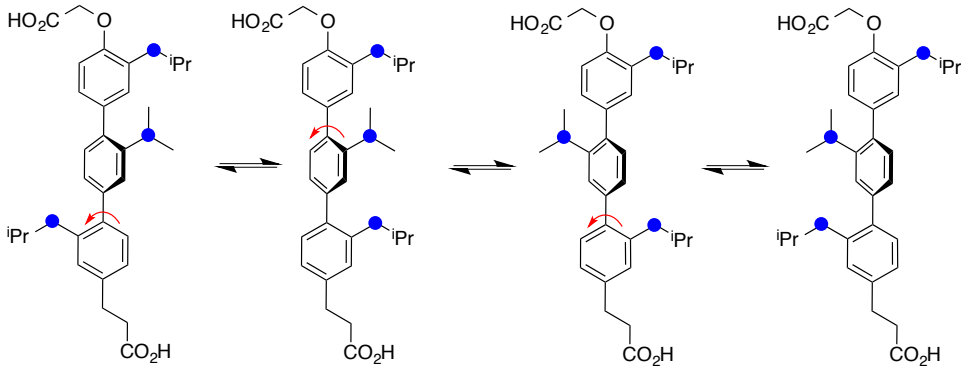


Figure 1.3. Equilibrating conformers of Hamilton's terphenyl minimalist mimic in solution.

1.4 Exploring Key Orientations (EKO) and Exploring Key Orientations on Secondary Structures (EKOS)

To fully explore the potential of a small molecular scaffold in targeting PPI interface motifs, we developed the *Exploring Key Orientations (EKO)* strategy. Briefly, in EKO each set of three residues on a protein at a crystallographically characterized PPI interface is categorized in terms of six Cartesian coordinates corresponding to the three sets of $C\alpha$ and $C\beta$ atoms. (Figure 1.4) Similarly, each kinetically and thermodynamically accessible conformation of a minimalist mimic of a PPI interface, *an interface mimic*, can be characterized in the same way, regardless of whether it resembles a secondary structure or not.

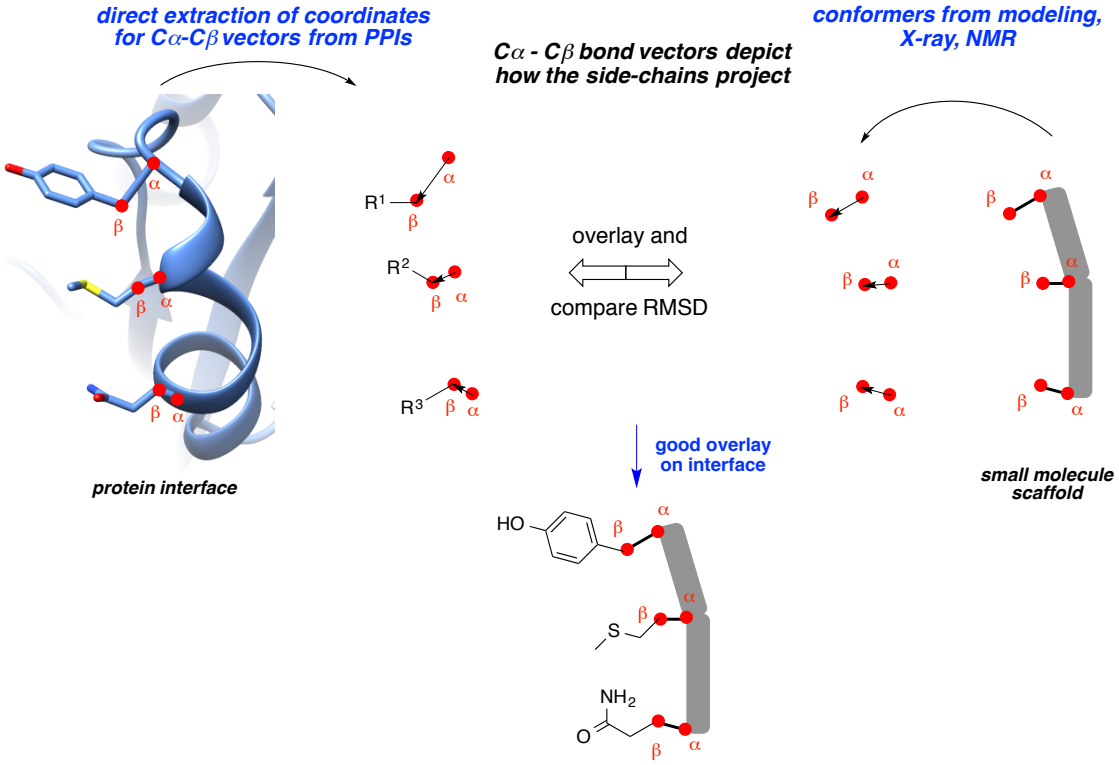


Figure 1.4. Identification of good interface mimics with Exploring Key Orientations.

EKO (Exploring Key Orientations)[58] and EKOS (Exploring Key Orientations on Secondary structures, described in Chapter VI)[57] are strategies to facilitate correlations of accessible solution state conformations of certain chemotypes with protein-protein interaction (PPI) interfaces, and with secondary structures, respectively. Specifically, EKO and EKOS are designed to work with chemotypes that involve semi-rigid organic scaffolds with three amino acid side-chains. EKO is designed to identify chemotypes of this kind that perturb PPIs. It involves molecular dynamics to generate a comprehensive set of accessible conformations of these molecules, characterization of each of these conformations in terms of the side-chain orientations that they express, then data mining

to match these with side-chain orientations found at PPI interfaces. The implication is that if the scaffold can present side-chains in the same orientation as an interface region involving one protein in a PPI, then it might be able to displace that protein, or at least perturb the interface.

1.5 Conclusion

Developing general and novel methods to find small molecules to target PPIs can be challenging but at the same time highly rewarding due to the importance of PPIs in biological systems. Minimalist mimics represent one of the new approaches and some interesting bioactive small molecules for PPIs have been discovered in this way. The novel computational approaches EKO and EKOS can provide useful theoretical predictions to complement the application of minimalist mimics. These computational models take advantage of the naturally existing binding pockets at the PPI interface regions and further development of these methods may lead to a general solution to the problem of finding suitable small molecule scaffolds for particular PPIs.

My PhD study is focused on the design and synthesis of several novel small molecule scaffolds as minimalist mimics for PPI targets. These small molecules are proven to be suitable scaffolds in the EKO and EKOS analysis and some of these molecules are biological active for selected PPI targets.

CHAPTER II

OLIGO-PIPERIDINE-PIPERIDINONES AS MULTI-FACETED SECONDARY STRUCTURE MIMICS*

2.1 Introduction

Minimalist mimics of secondary structures display amino acid side-chains on backbones more rigid than, and structurally different to, polypeptide chains.[42,47,59] Mimics that can be made conveniently with a variety of amino acid side-chains tend to be the most valuable, hence modular syntheses involving amino acid-derived starting materials are ideal.[45,60,61] Besides these synthetic considerations, it is also important to gain an understanding of how well ideal secondary structures can be represented by the best-fitting accessible conformer from the mimic's conformational ensemble. Such analyses indicate how a minimalist mimic can optimally fit on one secondary structure relative to another, and which side-chain spacings are best represented. This is valuable information because otherwise, for example, it is unclear which of the so-called ideal α-helical mimics actually fit better on other secondary structures, and which ones best present a given side-chain combination (*eg i*, *i*+3, *i*+4 or *i*, *i*+4, *i*+5).[57]

---

*Reprinted in part with permission from "A Multi-faceted Secondary Structure Mimic Based on Piperidine-piperidinones", Dongyue Xin, Lisa M. Perez, Thomas R. Ioerger and Kevin Burgess, *Angew. Chem. Int. Ed.* **2014**, *53*, 3594-3598. Copyright 2014 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim and from "Extended Piperidine-piperidinone Protein Interface Mimics", Dongyue Xin, Arjun Raghuraman and Kevin Burgess, *J. Org. Chem.* **2015***, 80*, 4450-4458. Copyright 2015 American Chemical Society.

The minimalist mimic oligo-pyrrolidine-pyrrolidinone (**A**, in Figure 2.1) was previously developed in Dr. Burgess's research group.[58,62,63] Conformational studies indicated that this scaffold exists in an extended conformation and it is able to mimic the β-strand and β-sheet regions of PPIs. Previous studies showed that this scaffold could be used to mimic the β-strand at the dimerization interface of HIV-1 protease.[58] Derivatives of scaffold **A** were shown to effectively promote the dissociation of HIV-1 dimer and inhibit its enzymatic activity.[58]

Scaffold **A** is a novel and bioactive minimalist mimic, but it also has some limitations. Firstly, accessibility of scaffold **A** is limited by the difficulties in its synthesis. **A** vulnerability in syntheses of substrates **A** is the enolizable, and therefore stereochemically delicate, *CH*.[63] The epimerization of the chiral centers on the scaffold will lead to ambiguity in its chemical structure and its potential interaction with PPI targets. Another problem in syntheses of scaffolds **A** is that fragments bearing bulkier amino acid side-chains tend to retard coupling reactions. Therefore, bulky amino acid side chains such as valine and threonine cannot be successfully incorporated into the scaffold. Besides, scaffold A is mainly designed to mimic the β-strand and β-sheet regions of PPIs, so it has limited application in PPI targets featuring helical secondary structures at the interface regions. Therefore, the design of novel oligomic mimic scaffolds that are synthetically accessible and conformationally distinctive from scaffold **A** is highly desirable. This chapter will discuss the design, synthesis and conformational studies for oligo-piperidine-piperidinone scaffold (**1**, Figure 2.1) as a novel oligomeric minimalist mimic.[64-66]

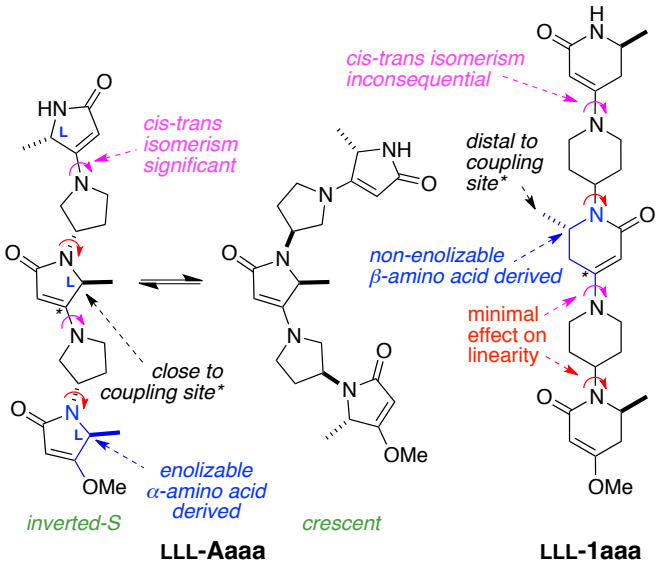


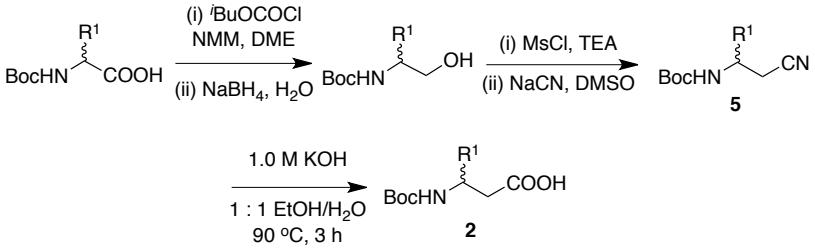
Figure 2.1. A comparison between oligo-pyrrolidine-pyrrolidinone and oligo-piperidine-piperidinone. Oligo-pyrrolidine-pyrrolidinone **A** can snake between various *S*- and crescent-shaped conformers, whereas all the inter-ring σ-rotations in **1** together represent only one significant degree of freedom making the scaffold overall more linear.

2.2 The Design and Syntheses of Novel Oligo-Piperidine-Piperidinones as Secondary Structure Mimics

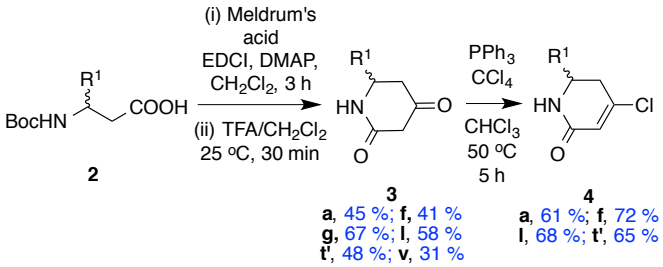
To solve some of the problems faced by scaffold A and target different extended PPI regions, the oligo-piperidine-piperidinone minimalist mimic **1** was designed. An attribute of scaffolds **1** is that they are based on β-amino acids where a methylene group insulates the chiral center from stereomutation. In addition, in comparison with scaffold **A**, we postulated this steric effect would be less pronounced in construction of **1** where the side-chains are one methylene removed from the electrophilic coupling site. Therefore, mimic **1** is more synthetically accessible than **A**.

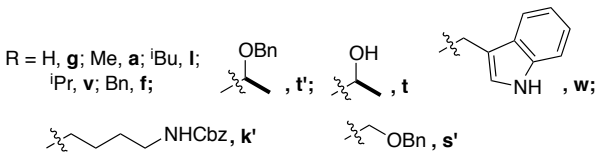
Minimalist mimics **1** can prepared from amino acid-derived starting materials with a modular synthesis procedure. Syntheses of molecules **1** began with preparation of β-substituted β-amino acid derivatives **2** using a novel combination of literature procedures[67,68] that allowed us to obtain multi-gram amounts, without chromatography in most cases (Scheme 2.1). These amino acids **2** were reacted with Meldrum's acid, a known reaction for α-amino acids,[69] to form six-membered homologs of tetramic acids (Scheme 2.2). Removal of the Boc-protecting group gave the piperidinediones **3** that were converted to the vinylic chlorides **4**. Those analogs with threonine side-chains have an inherent marker for loss of stereochemical fidelity; as anticipated, no epimerization was observed, at least in the syntheses of the compounds containing Thr. Throughout this thesis, compounds are numbered according to the scaffold with lower case one-letter codes relating their side-chains to the parent amino acids (*eg* **f** for Phe, **s** for Ser, **s'** for the $-CH_2O^tBu$, **t** for -CH(OH)Me, and **t'** for -CH(OBn)Me; named from *N*- to *C*-terminus just as in peptides).

Scheme 2.1. Syntheses of the β-amino acid derivatives.



Scheme 2.2. Syntheses of the electrophilic *N*-caps **4**.

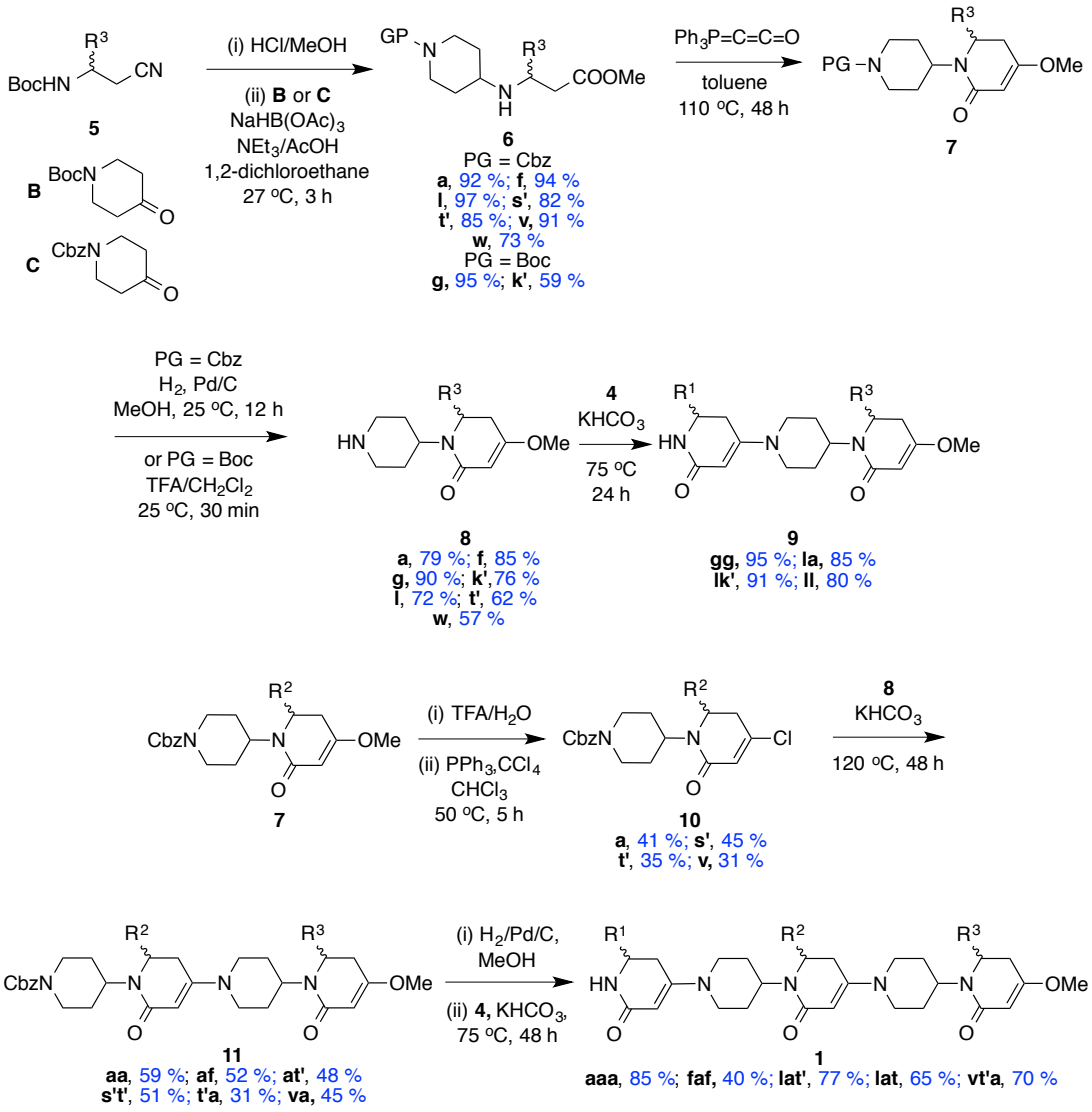




Nitriles **5**, intermediates in the syntheses of the β-amino acids **2**, were simultaneously *N*-deprotected and hydrolyzed, then reductively coupled with the known, and commercially available, synthons **B**[70] or **C**[71] to give the amines **6** (Scheme 2.3). Reaction of these β-amino esters with Bestmann's ylide[72] gave the protected intermediates **7**. Compounds **5** to **7** were isolated, without chromatography, on multigram scales. *N*-Deprotection of the intermediates **7** gave the nucleophiles **8**. Amines **8** were then condensed with the electrophiles **4** to give the scaffolds **9** bearing two side-chains. As anticipated, this process seems largely unaffected by steric demands of the side-chains since the coupling yields were uniformly high.

In a divergent step, intermediates **7** were *C*-deprotected, then converted to the vinylogous chlorides **10**. Convergence was then possible via coupling of the nucleophiles **8** with the electrophiles **10** to give the extended systems **11** that were then *N*-deprotected and capped with the electrophiles **4**.

Scheme 2.3. Syntheses of the target materials **1**.



5

B

C

6
PG = Cbz
a, 92 %; **f**, 94 %
l, 97 %; **s'**, 82 %
t', 85 %; **v,** 91 %
w, 73 %
PG = Boc
g, 95 %; **k'**, 59 %

7

8
a, 79 %; **f**, 85 %
g, 90 %; **k'**,76 %
l, 72 %; **t'**, 62 %
w, 57 %

9
gg, 95 %; **la,** 85 %
lk', 91 %; **ll**, 80 %

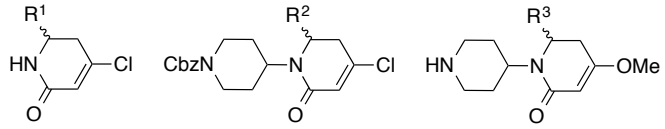
7

10
a, 41 %; **s'**, 45 %
t', 35 %; **v,** 31 %

11
aa, 59 %; **af**, 52 %; **at'**, 48 %
s't', 51 %; **t'a**, 31 %; **va,** 45 %

1
aaa, 85 %; **faf,** 40 %; **lat'**, 77 %; **lat**, 65 %; **vt'a**, 70 %

The syntheses of mimics **1** may seem complicated, but they are driven by the simple modularity concept expressed in Figure 2.2. Overall, the divergent-convergent syntheses of compounds **1** pivot around synthons **7**; these were converted to nucleo- and electrophiles that were joined to elongate the scaffold. Fragments **8** and **10** in Scheme 2.3 are similar to *C*- and *N*-protected amino acids in peptide syntheses.

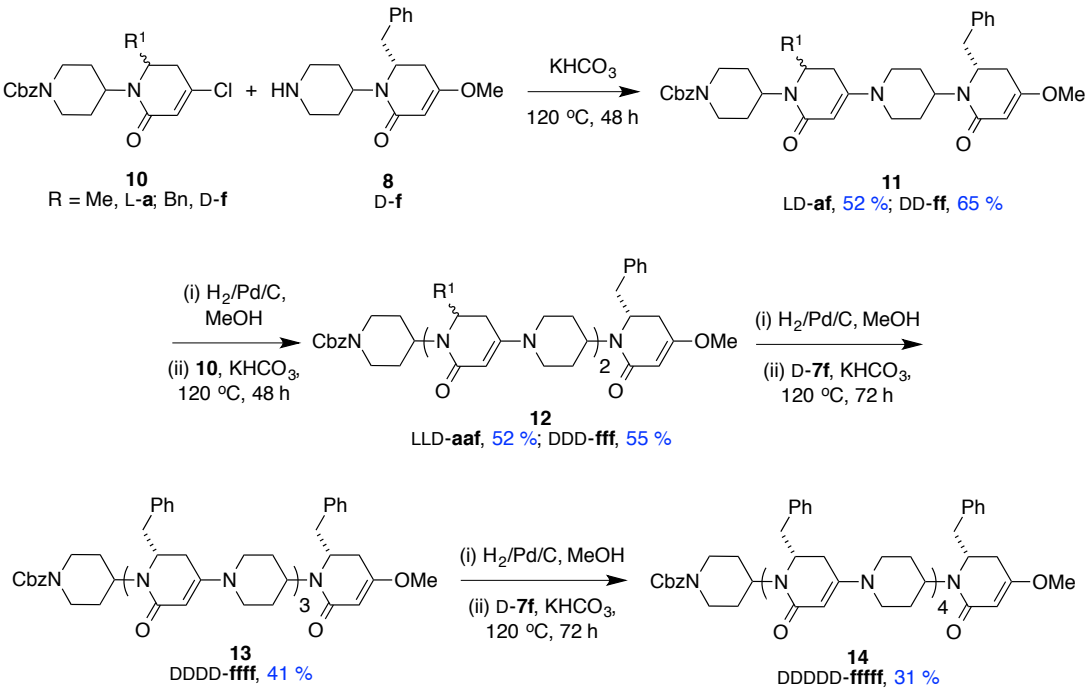


N-cap **4** | *N-protected chain growth intermediate* **10** | *C-protected chain growth intermediate* **8**

Figure 2.2. *N*-Capping, *N*-protected, and *C*-protected synthons like **4**, **10**, and **8**, are pivotal in divergent-convergent routes to scaffolds **1**.

Scheme 2.4 illustrates how iterative couplings under elevated temperature and mildly basic conditions were suitable for synthesis of **14fffff** with five repeating units derived from Phe.[66] The synthesis indicates that the coupling reactions are efficient enough to enable the synthesis of long oligomers. Figure 2.3 shows the MALDI MS of this product has a single ion at the predicted m/z of 1507.8 for M+H (plus 1529.8 for M+Na, and 1545.8 for M+K). This product is considerably less polar than typical peptides, being soluble in dichloromethane, methanol, and DMSO.

Scheme 2.4. Synthesis of the piperidinones-piperidine mimic **14fffff** with ten linked rings.



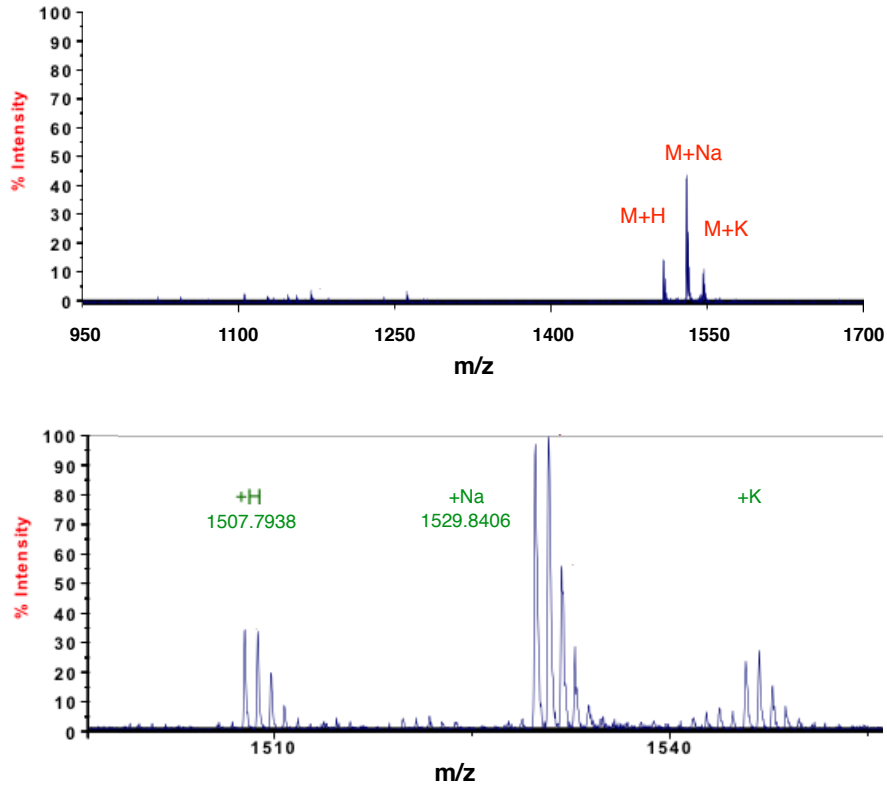


Figure 2.3. HRMS (MALDI+) spectrum of DDDDD-**14fffff**.

2.3 Conformational Studies of Mimic 1

No spectroscopic technique can characterize conformational ensembles of small molecules like **1** or **A** in solution because they equilibrate between hundreds solution states that have significantly different side-chain orientations. Techniques like CD and NOESY/ROESY NMR give a conformational average, that does not correspond to any real state.[45] An X-ray structure shows whatever mimic conformation happened to crystallize, where part of the shape is governed by crystal packing forces. Those packing forces are significant enough that some solid-state conformations may not be preferred in solution. Indeed, a solid state mimic conformation that does *not* resemble

the target secondary structure has been reported by Hamilton *et al*,[55] and this type of occurrence should not be surprising because crystal packing forces are relatively large.

Three intermediates (**9gg**, **11af, 1faf**) were isolated[65] and crystallized in the course of the synthetic studies described above. Representations from X-ray crystal structures of those compounds are shown in Figure 2.4.

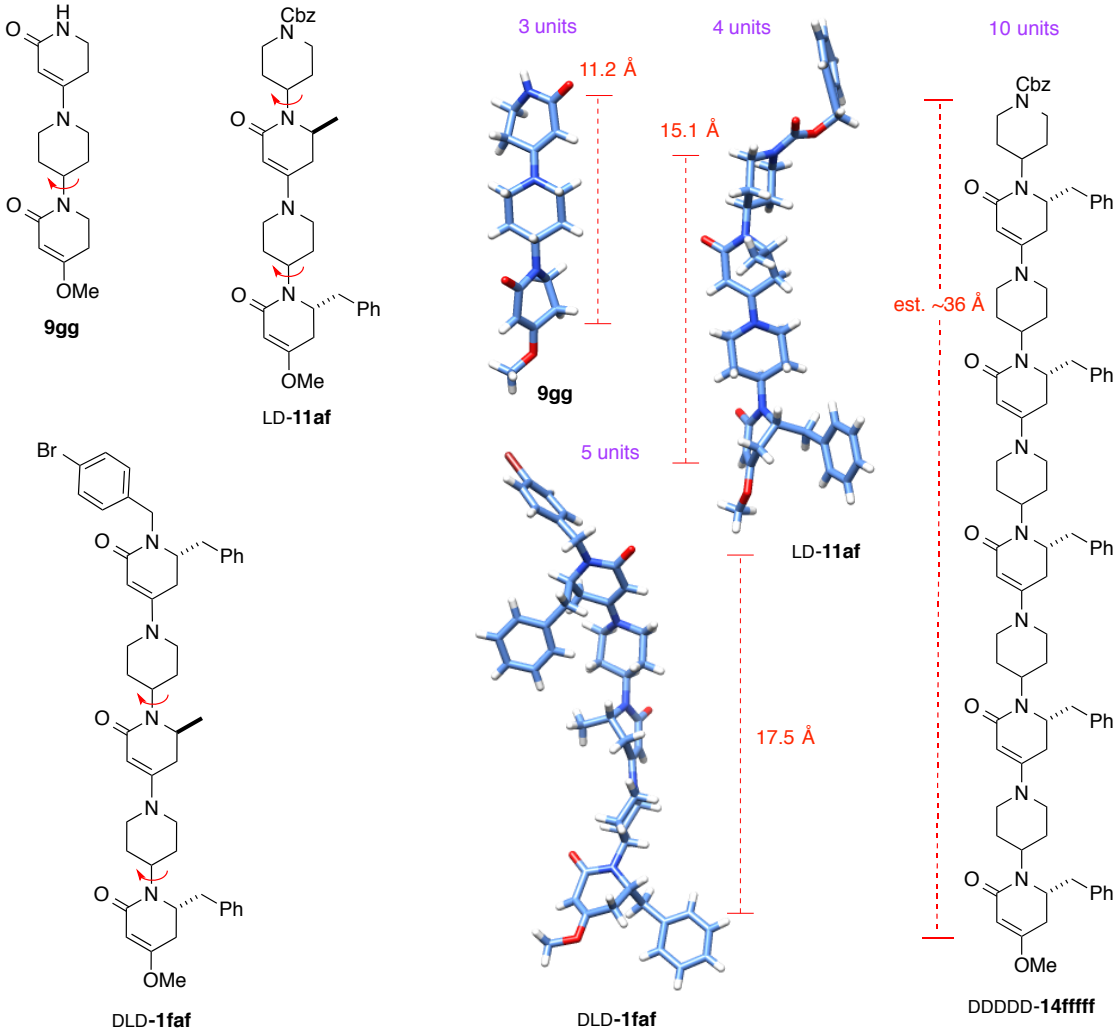


Figure 2.4. X-ray crystal structures of scaffolds **9gg**, **11af** and **1faf**. The *N*- to *C*-dimension of **14fffff** was estimated to be around 36 Å.

Solid state structures of **9gg**, **11af**, and **1faf** reveal these molecules crystallize as conformers that have different dihedral angles for the linkages marked with red arrows in Figure 2.4. Rotation around bonds associated with these red arrows has a profound effect on the side-chain orientations. This supports the assertion that minimalist mimics have the potential to adopt conformers that can overlay on more than one ideal secondary structure or interface region. Compounds **9gg**, **11af**, and **1faf** have 3, 4, and 5 piperidinone/piperidine rings, and they measure (*N*- to *C*-termini) 11.2, 15.1, and 17.5 Å, respectively. Those dimensions correspond to 3.73, 3.77, and 3.50 Å respectively per piperidinone or piperidine ring, and the difference between these figures (0.27 Å) reflects the extent of kinking or curvature in the overall conformation. None of these solid state conformations overlay very well on ideal secondary structures, but within this series **9gg** and **11af** are the most extended (*cf* sheet structures), and **1faf** is most helical. Compound DDDDD-**14fffff** was *not* crystalized, but the solid state data suggests this compound might have *N*- to *C*-dimensions of (10 x 3.50) to (10 x 3.77), or around 36 Å in the solid state.

A problem that has not been addressed in studies of minimalist mimics is how to evaluate the relevance of a solid-state structure to a conformational ensemble in solution. Here we suggest EKOS (described in Chapter VI) can be used to determine if a solid-state conformation from an X-ray structure is present in the simulated conformational ensemble and, if it is, provide an estimate of the energy difference between the simulated solution-state conformer closest to the solid-state structure and the lowest energy one overall. EKOS also makes it possible to overlay the single conformer represented in a

solid-state structure on all ideal secondary structures to find the one it best matches and the preferred side-chain combination. This information is generated free from the bias of the researcher(s) who may have designed the scaffold to mimic a particular secondary structure and side-chain conformation.

An example of the use of EKOS as described above is as follows. Figure 2.5a shows data for a solid-state structure of DLD-**1faf**. That structure overlaid on an ideal α-helix with an RMSD (root-mean-square deviation) of 1.5 Å (based on the six *C*α + *C*β coordinates of the side-chains; Figure 2.5b). To estimate the energy of the conformer represented by the X-ray structure, *but in solution*, the EKOS was used to systematically overlay the simulated conformational ensemble on the X-ray structure to find the conformer with the best fit. In fact, that conformer (matched to 0.46 Å RMSD; Figure 2.5c) had a simulated energy of 0.91 kcal•mol^{-1} relative to the lowest one generated, *ie* it is predicted to be significantly populated in solution. The simulated conformer that overlaid most closely on the ideal α-helix did so with a similar RMSD and energy (0.50 Å, 0.84 kcal•mol^{-1}, Figure 2.5d). Overall, these data indicate there is a conformer in the simulated ensemble that represents an ideal a-helix better than the one that crystallized. Moreover, the simulated solution-state conformers closest to the solid state structure and to an ideal a-helix seem to have about the same relative energies. The conformer from the X-ray and the most α-helical one from the simulation both had the same side-chain correspondence: i, i+5, i+10.

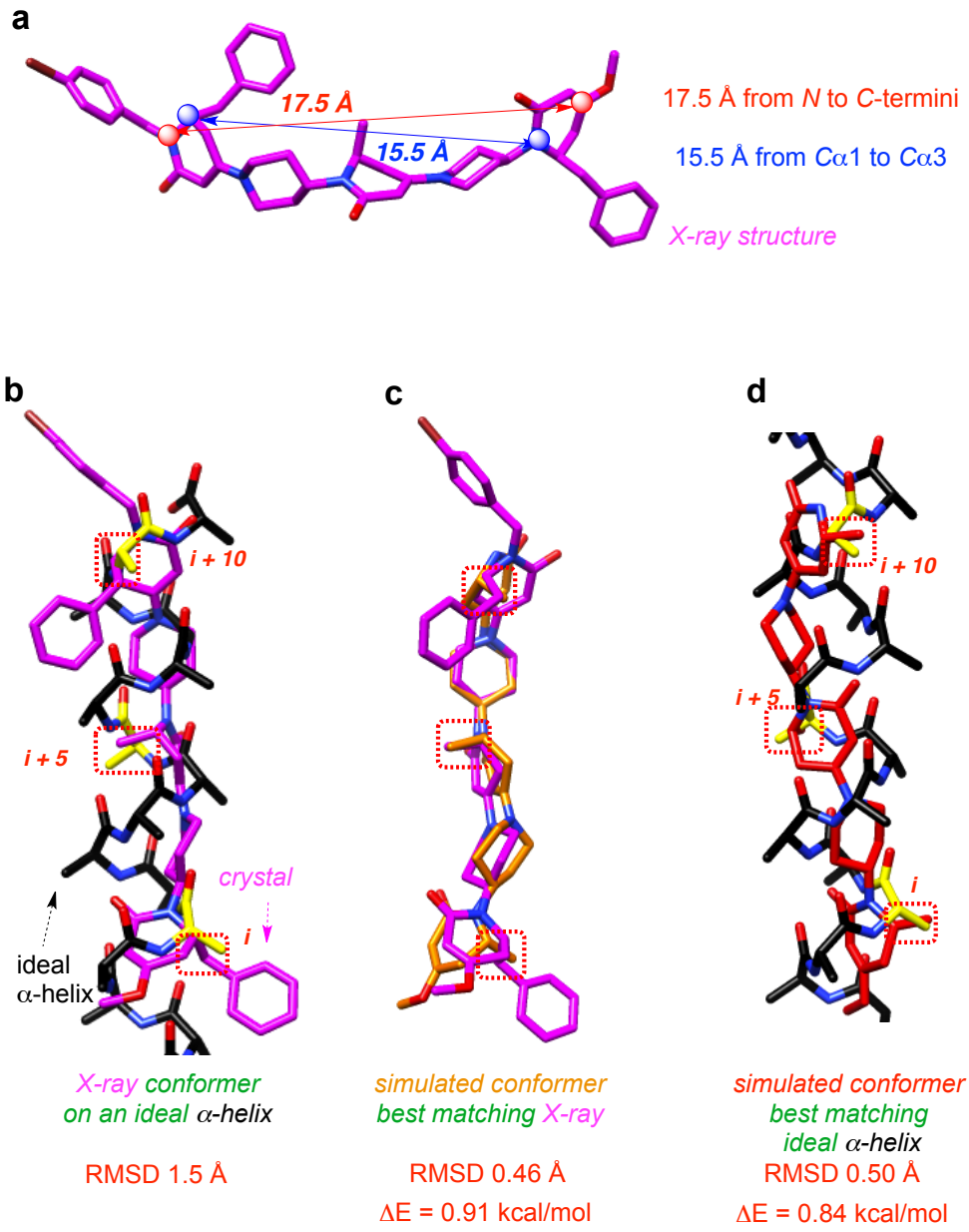


Figure 2.5. Analysis of the X-Ray crystal structure of DLD-**1faf**. **a** X-Ray crystal structure of DLD-**1faf**. That structure is shown overlaid upon: **b** an ideal α-helix; and, **c** the simulated conformer best matching the X-ray structure. **d** The simulated conformer best matching an ideal α-helix is overlaid on an ideal α-helix. **e** Two copies of the X-ray structure overlapping on the *C*- and *N*-termini illustrates the helical structure is anticipated to repeat every 3 nm and 5 residues.

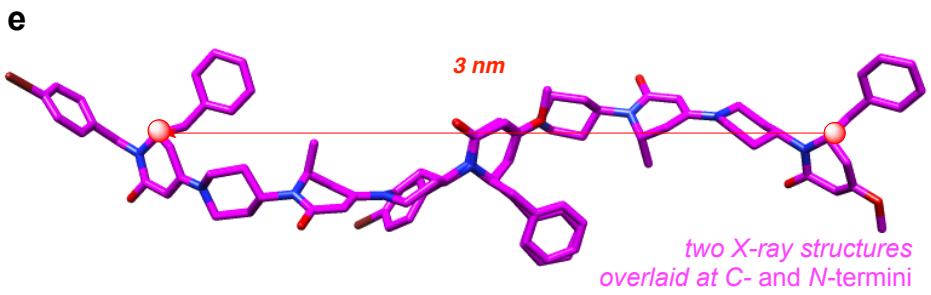


Figure 2.5. Continued.

Figure 2.5e shows copies of the X-ray structure arranged such that the *C*-terminal piperidinone of one is superimposed on the *N*-terminal piperidinone of the other. This illustrates how the solid-state conformer is anticipated to coil in a helical way, repeating every 5 residues and 3 nm. Parenthetically, Whitesides and co-workers communicated repeated reductive couplings of unfunctionalized 4-piperidinones to give "linker rods".[73] A "rod" with four piperidine repeats, was 1.6 nm long in the solid state (X-ray), comparable to 1.7 nm for **1faf** which has *five* rings, Figure 2.5a. Whiteside's structure is like a straight rod whereas **1** twists to accommodate the chiral piperidinones making it shorter between the *N*- and *C*-termini.

CD spectra of minimalist mimics are rarely informative because it is unusual to have a series that represents systematic changes to the scaffold structure.[74] However, in this particular study, several oligo(piperidinone-piperidine)s of different lengths were available, and this facilitated meaningful correlations of elipticities with structure. Figure 2.6a shows UV spectra of compounds represented by the generic structure **B** (n = 1 is **11ff**; n = 2, **12fff**; n = 3, **13ffff**; and, n = 4, **14fffff**; D-configuration throughout). Absorbance maxima at approximately 290 – 310 nm (MeOH) are attributed to additive

effects of the enamide chromophores (red box on **B**); more absorbance correlates with higher values of *n*. If the compounds were non-randomly distributed amongst stereoisomeric secondary structures (*eg* right- and left-handed helices) then increased molar elipticities would be expected as *n* increases. Figure 2.6b shows that molar elipticities in the 290 – 310 nm wavelength region do, in fact, increase with *n*, implying some degree of ordering in solution.

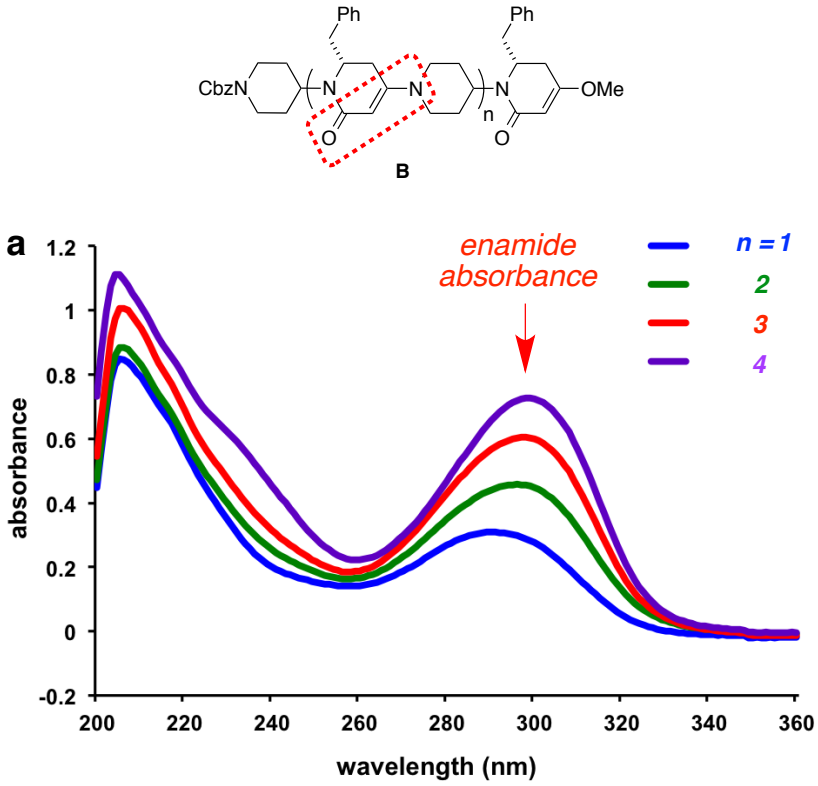


Figure 2.6. UV and CD spectra of oligo-piperidinone-piperidines. **a** UV; and, **b** CD spectra of **B** (*n* = 1 is **11ff**; *n* = 2, **12fff**; *n* = 3, **13ffff**; and, *n* = 4, **14fffff**) in MeOH. UV spectra were recorded at 20 μM, and CD at 200 μM.

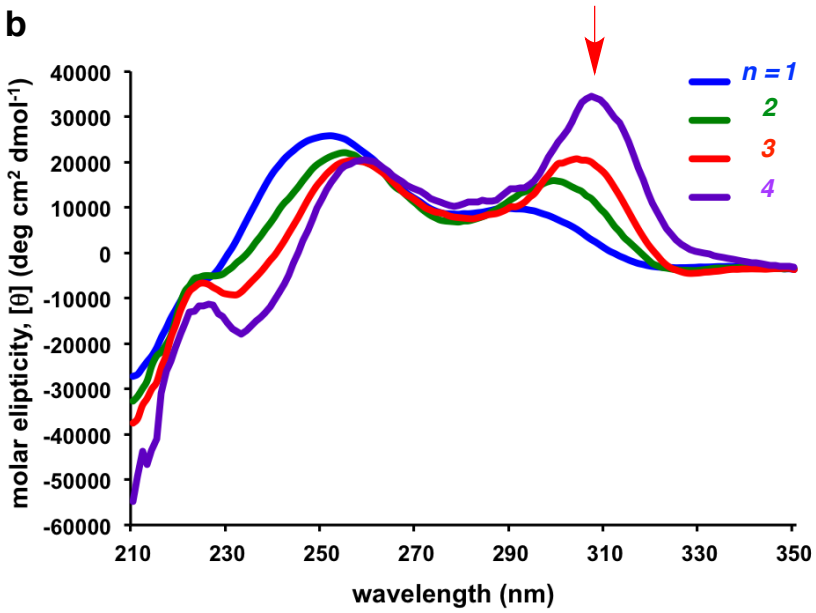


Figure 2.6. Continued.

We hypothesized scaffold **1** could present three amino acid side-chains in orientations resembling extended secondary structures with wide side-chain separations. To test our hypothesis, a large number of accessible (<3 kcal•mol^{-1}) conformations for each stereomer of **1** were simulated with quenched molecular dynamics,[75,76] then systematically matched with the most common *ideal* secondary structure elements with EKOS.[57]

Figure 2.7 features data for all stereomers of **1**; it shows the RMSDs for the best fitting conformers corresponding to each element of secondary structure. These data are calibrated relative to the average RMSDs for all the best fitting conformers for a given stereomer. Thus for LLL-**1aaa** 0.65 Å was the average for the seven best fitting conformers on the seven secondary structures indicated. This graphic indicates the LLL-isomer is predicted to mimic a sheet-turn-sheet closely, ideal helical structures less well,

and extended parallel β-sheet- and -strand structures very poorly (Figure 2.7a). Comparing this with the plot for the DLD-isomer (selected because the X-ray data in Figure 2.5a was accumulated for that isomer) shows a similar tendency towards helical conformations. The fact that the overlays on sheet-turn-sheet and antiparallel β-sheet structures have identical RMSD values is indicative of the fact that the antiparallel β-sheet region is part of the sheet-turn-sheet motif, and this is where the mimic overlaid.

Figure 2.7i shows the corresponding data for LLL-**Aaaa**. Scaffold **A** has a strong tendency to mimic β-sheet structures, but not helical ones, and that trend is in fact consistent for all the stereomers of **A**.

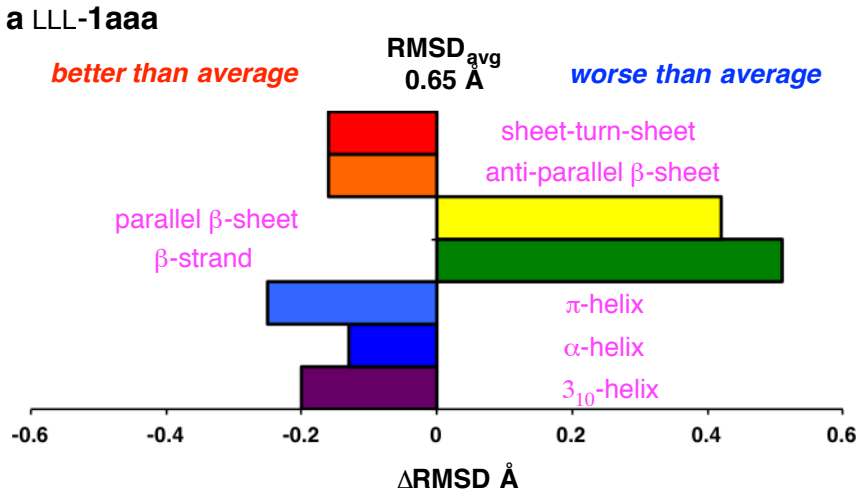


Figure 2.7. Predicted mimicry of secondary structures for: **a-h** all diastereomers of **1aaa** and **i** LLL-**Aaaa**.

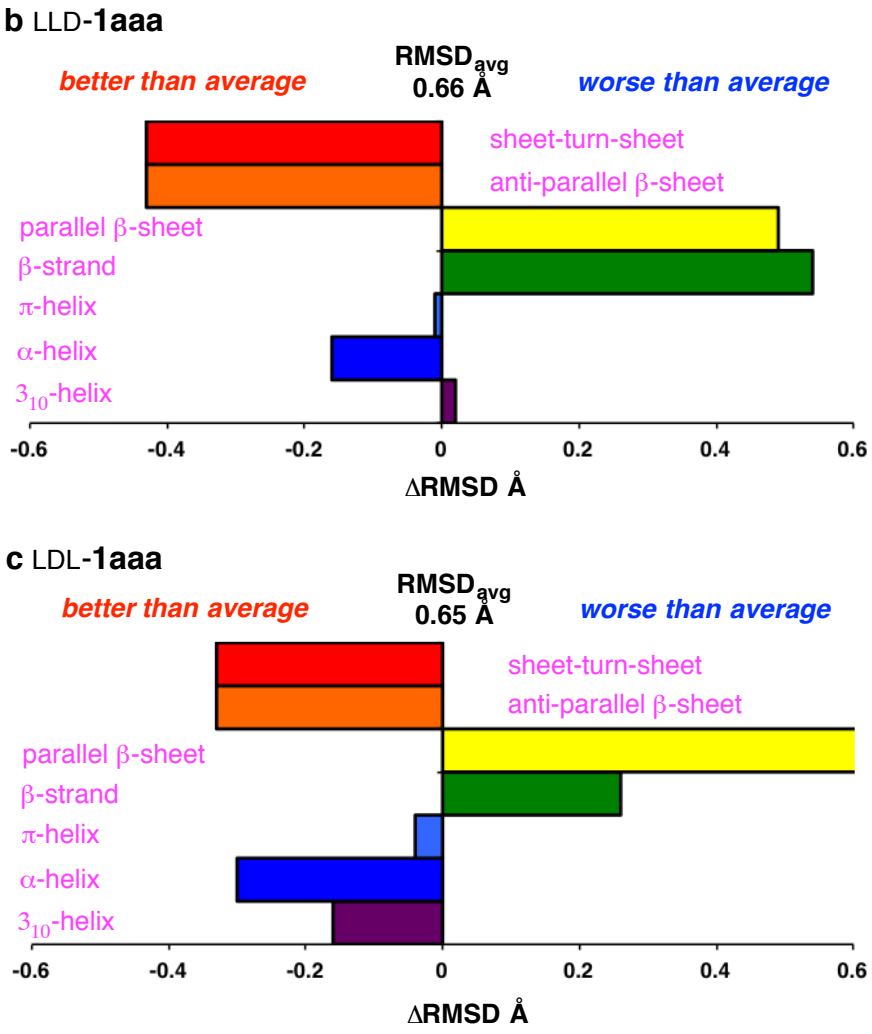


Figure 2.7. Continued.

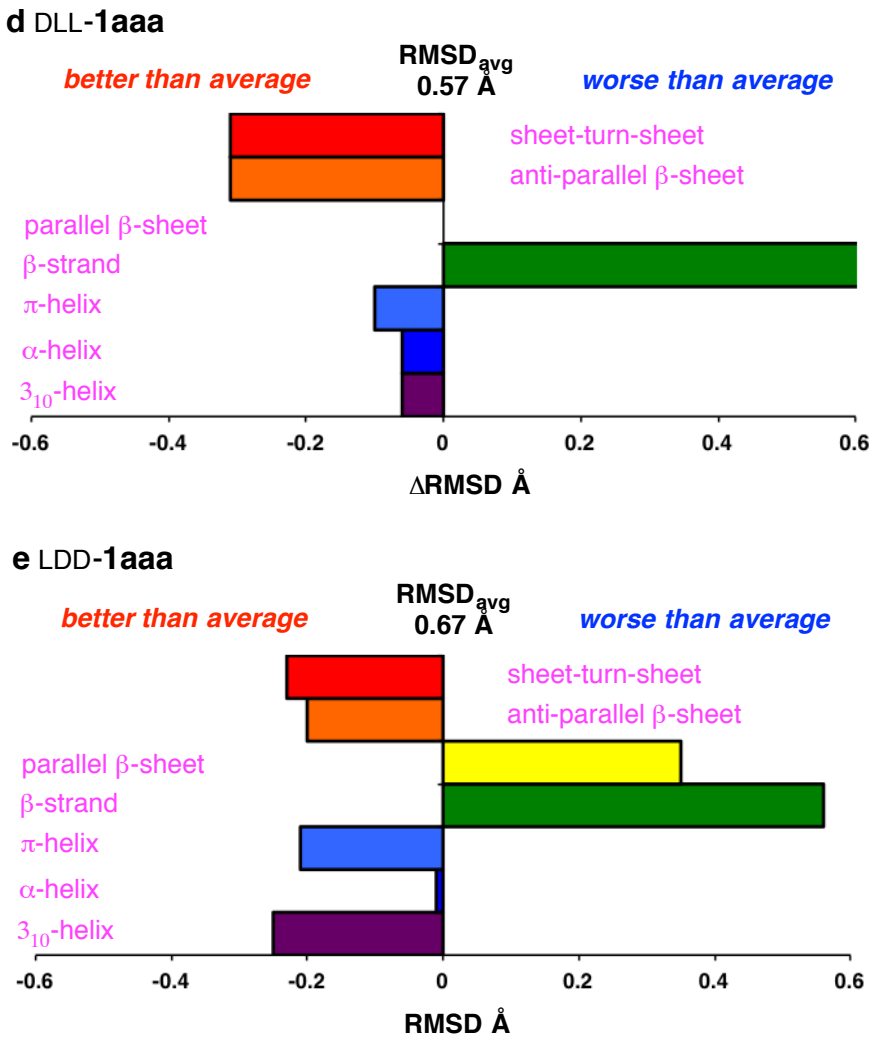


Figure 2.7. Continued.

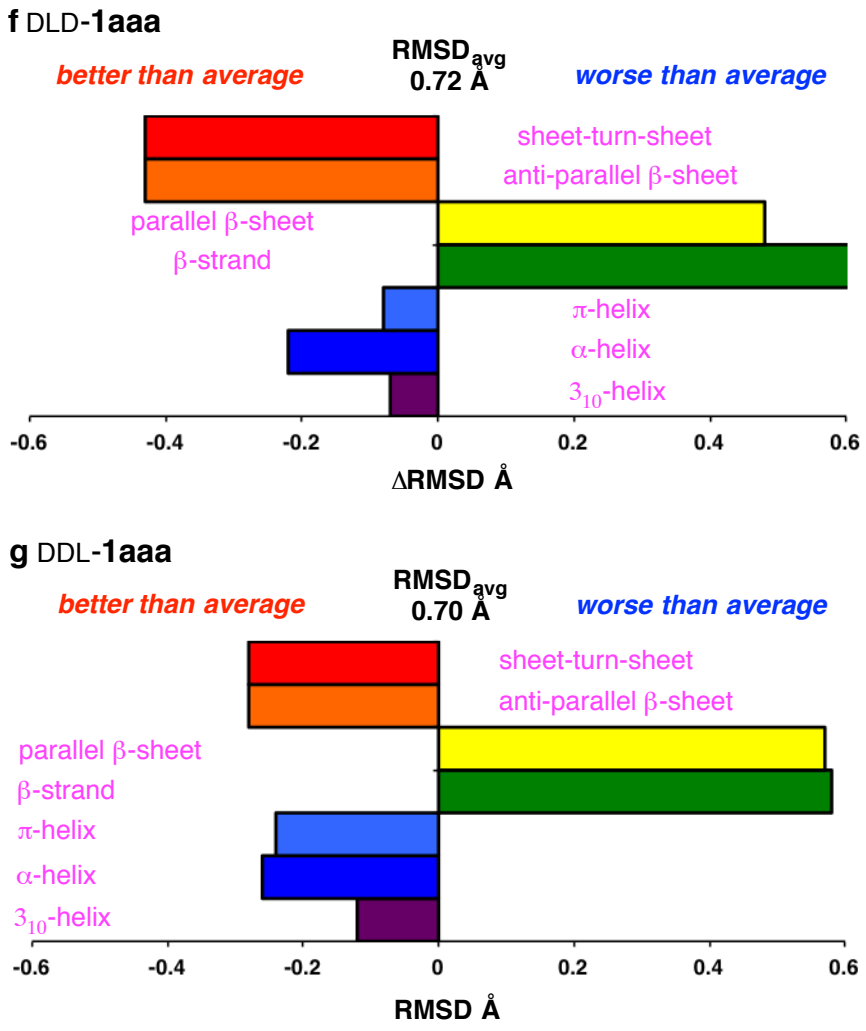


Figure 2.7. Continued.

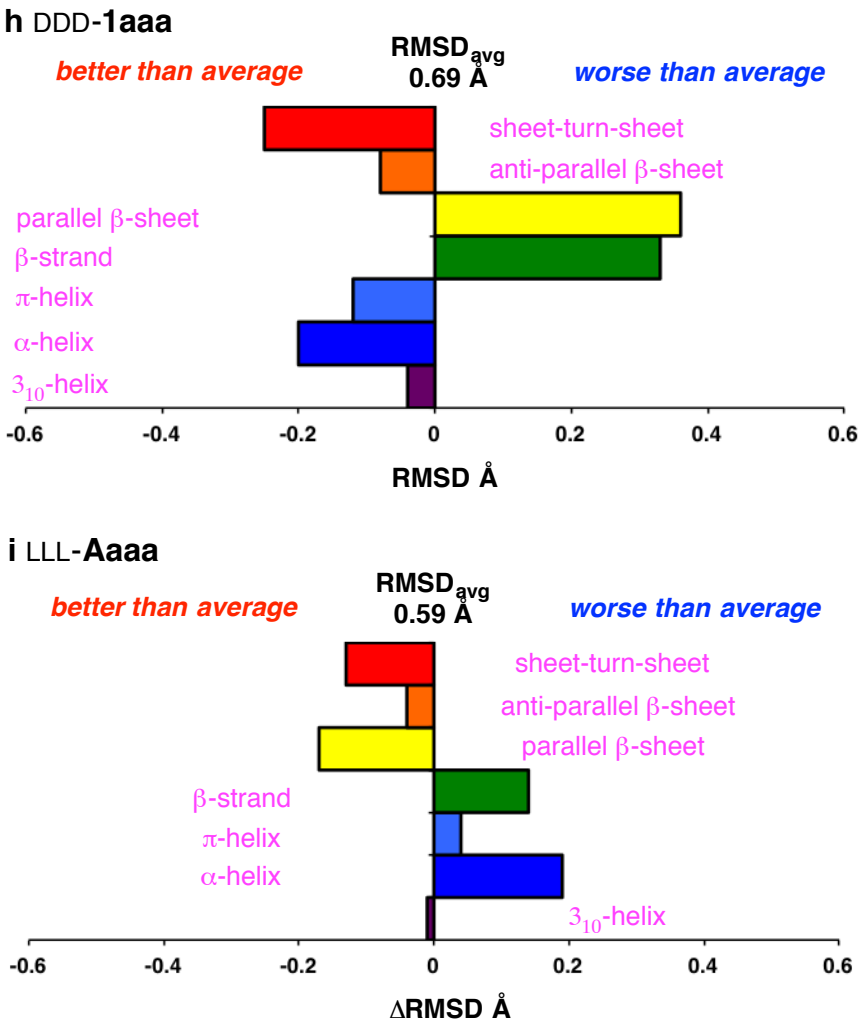


Figure 2.7. Continued.

2.4 Data-mining of All Crystallized PPIs in Protein Data Bank

Minimalist mimics are usually designed to perturb real protein-protein interactions (PPIs) by displacing a secondary structure at the interface. However, interfaces rarely have *ideal* secondary structures, and many do not feature secondary structures at all. Exploring Key Orientations (EKO) is data mining to compare

simulated conformational ensembles of minimalist mimics with crystallographically characterized PPI regions, irrespective of the secondary structures involved.[58] Here EKO analyses for LLL-**1aaa** and LLL-**Aaaa** were performed to determine if the biases for *ideal* secondary structures predicted above (using EKOS) are consistent with the statistical match of accessible conformations on *real* interface secondary structures. To do this we developed a script to assess the secondary structure content of the interface regions that best match simulated conformers. After analyses of tens of thousands of PPI structures, EKO found 68 interface regions that matched a simulated conformer of **1** with RMSDs of <0.3 Å, and 32 that similarly matched conformers of **A**. Figure 5 shows only LLL-**1aaa** matched on helical regions at PPI interfaces though it found relatively more sheet interface regions. LLL-**1aaa** and LLL-**Aaaa** had a similar proclivity to match with contiguous or non-contiguous motifs of no particular secondary structure (called "single segment" and "multiple segments" in Figure 2.8). Overall, these observations are consistent with those from comparisons of the simulated conformational ensembles with *ideal* secondary structures in Figure 2.7. Four representative overlays for LLL-**1aaa** at interfaces are illustrated in Figure 2.9.

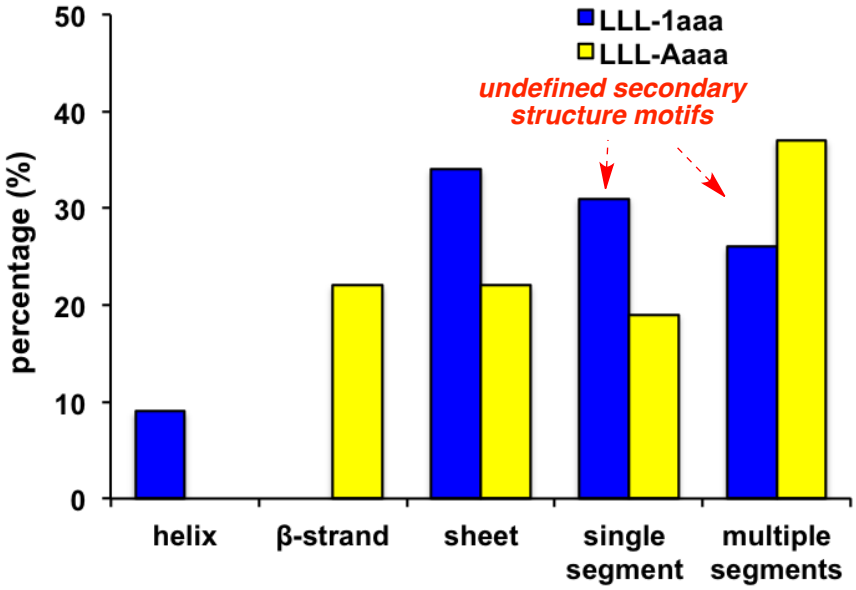


Figure 2.8. The most favorable PPI matching regions from EKO analyses of LLL-**1aaa** and LLL-**Aaaa** compared in terms of the secondary structure motifs that they overlay upon. "Single segment" means the side-chains are three residues on a single chain segment, whereas "multiple segments" refers to situations for which the three important side-chains are on non-contiguous regions of the protein.

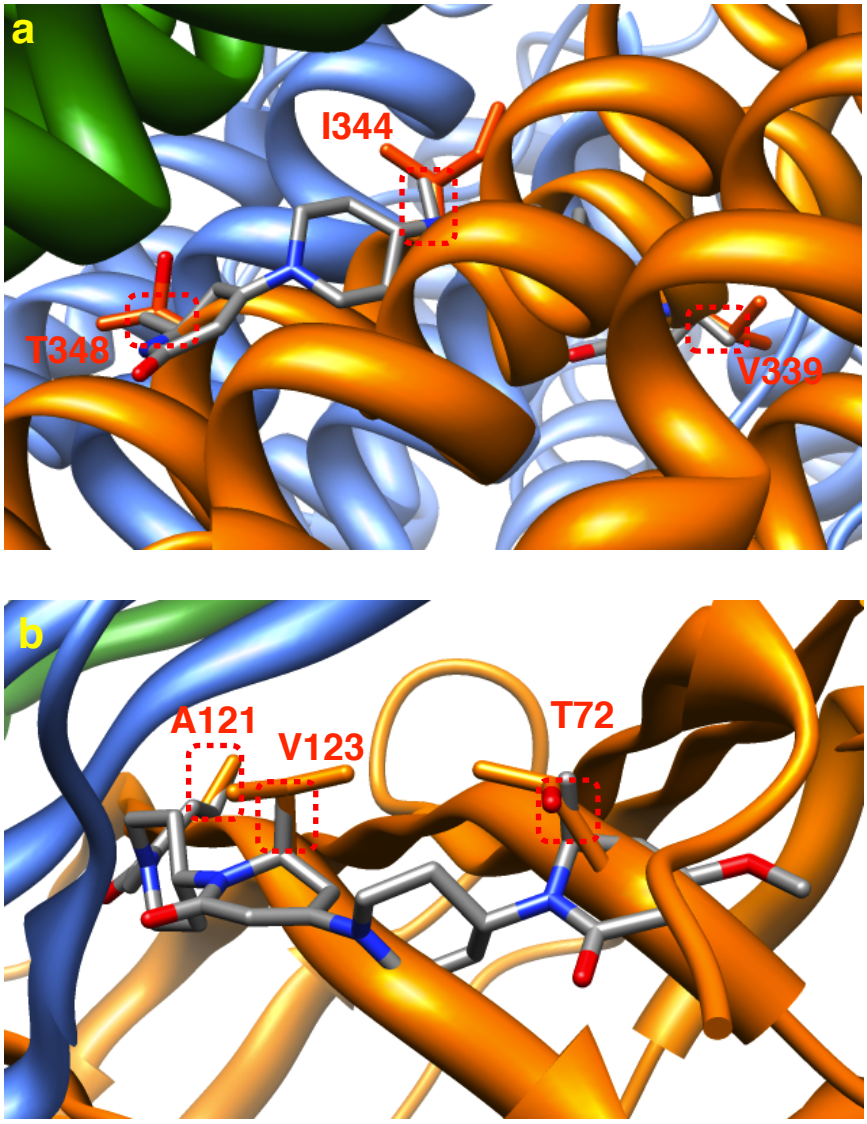


Figure 2.9. EKO predicts LLL-**1aaa** can overlay well on different secondary structure motifs at PPI interfaces. **a** Helical motif of L-aspartate ammonia lyase, RMSD = 0.29 Å, PDBID: 1jsw. **b** β-Sheet of ectodysplasin A, RMSD = 0.27 Å, PDBID: 1rj7. **c** On a contiguous chain-segment without a defined secondary structure for prolyl-^{t}RNA synthetase, RMSD = 0.21 Å, PDBID: 1h4s. **d** On discontiguous segments of a PPI interface for RNA polymerase II, RMSD = 0.30 Å, PDBID: 2cw0.

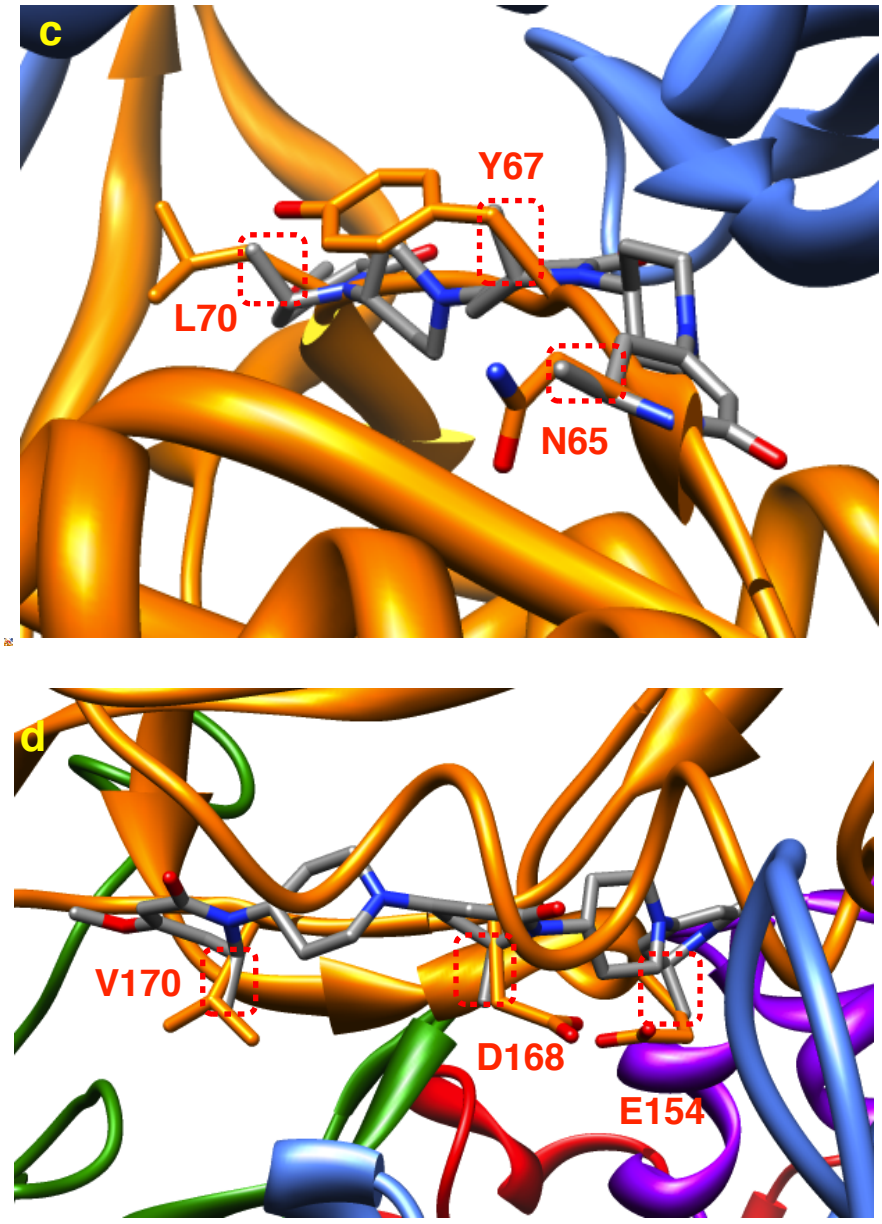


Figure 2.9. Continued

2.5 Conclusions

In summary, a route to mimics **1** that is not vulnerable to epimerization, and involves facile couplings, was devised. Solution phase syntheses of mono-, di-, tri-,

tetra-, and penta-oligo(piperidinone-piperidine)s were convenient. Solid state structures of three compounds of intermediate lengths in this series showed they tend to adopt extended conformations, and provided evidence to suggest a 36 Å longitudinal dimension for the pentamer **14fffff**. CD studies of milestone intermediates in this synthesis indicates the conformational ensemble tends towards progressively more ordered structures in solution.

Only six methylene groups differentiate the structures of **1aaa** and **Aaaa** but that drastically changes side-chain orientations in their solution conformers. Overall, we assert that even though none of the conclusions about the secondary structure preferences for **1** in solution can be confirmed spectroscopically, and limited inferences can be drawn from X-ray data, it is much better to have insight from EKOS and EKO than to have none at all. EKOS indicated there are significant differences between the ways simulated preferred conformers of the featured scaffolds overlay on secondary structures; **1aaa** is predicted to be more linear, and has a bias towards helical motifs whereas **Aaaa** best mimics β-strands. Overall, simulated preferred conformers of molecules **1** and **A** optimally superimpose on different interface regions and proteins when they are systematically overlaid on a huge database of PPI structures using EKO; in other words, these two scaffold designs complement each other. Applications of mimics **1** in perturbation of a particular PPI based on EKO analyses will be discussed in Chapter IV.

CHAPTER III

THE DESIGN AND SYNTHESIS OF OLIGO-PIPERIDINE-PYRROLIDINONES*

3.1 Introduction

Previous studies from our laboratories have reported on the minimalist mimics **A** (Figure 3.1).[62,63] We were intrigued by the fact that small changes to the backbone structure **A** *must* affect the accessible conformations of these molecules, and therefore the types of protein-ligand interface regions they can resemble.[65] However, if the compounds are not synthetically accessible, then it will be impossible to test their effects in a cellular environment. Our group has shown mimics **A** can be prepared with a variety of amino acid side-chains, but not without limitations.[62,63] An alternative mimic **1** was developed in Chapter II and this scaffold is synthetically more accessible and conformationally distinctive compared with mimic **A**. Research described in this chapter covers a brief investigation of the pyrrolinones-piperidine mimics **15** that features a hybrid structure of **1** and **A**, and the discovery of a general and chemoselective methodology to synthesize mimics 15 and other β-enamino derivatives.

---

*Reprinted in part with permission from "Extended Piperidine-piperidinone Protein Interface Mimics", Dongyue Xin, Arjun Raghuraman and Kevin Burgess, *J. Org. Chem.* **2015**, *80*, 4450-4458. Copyright 2015 American Chemical Society and from "A Chemoselective Route to β-Enamino Esters and Thioesters", Dongyue Xin and Kevin Burgess, *Org. Lett.* **2014**, *16*, 2108-2110. Copyright 2014 American Chemical Society.

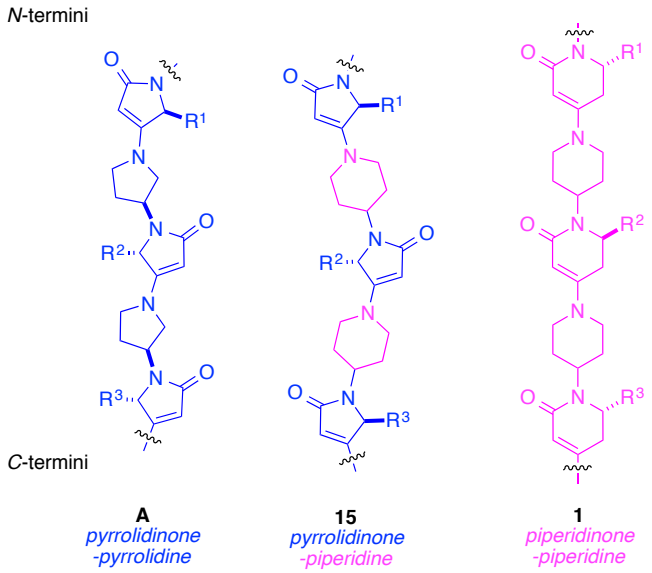


Figure 3.1. Structures of mimics **1** and **15**, conceptually derived from **A**.

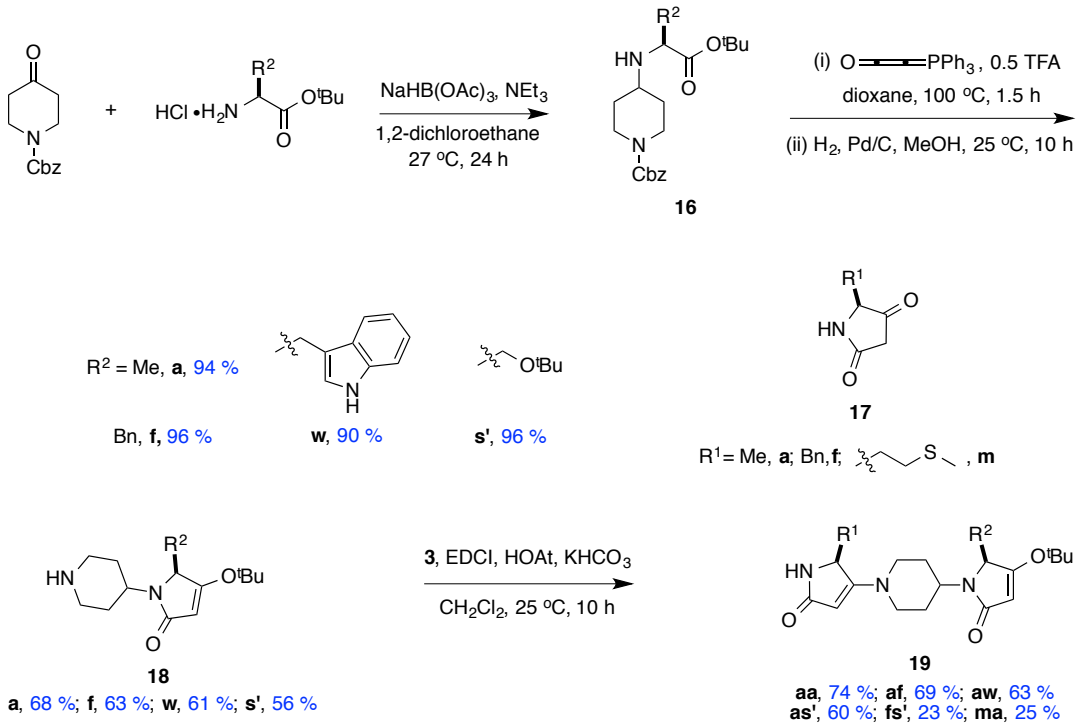
Scaffolds **A, 1** and **15**[66] have different conformational ensembles that orientate the three side-chains on each scaffold in different ways. Each scaffold will therefore overlay on *different* types of PPI interfaces, as determined using massive screening of the PDB with the data mining routine we developed as part of the EKO strategy.

3.2 Exploratory Studies on the Syntheses of Pyrrolidinone-piperidine Oligomers 15

Synthesis of the chiral center pyrrolidines **A** necessitated S_N2 displacement of a triflate from an enantiomerically enriched, *N*-protected, 3-hydroxypyrrolidine; this was not a reaction that proceeded cleanly from starting material to product.[62,63] Superficially, syntheses of mimics **15** appeared to be more facile because reductive amination to give the diamines **16** does not introduce a chiral center. In the event, diamines **16** *were* easy to prepare, and they were easily converted to piperidine-functionalized tetramic acid

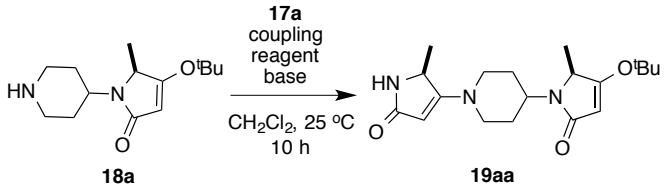
esters **18**;[77-79] the first chromatographic separation was required only at this stage, hence these monomers were conveniently prepared in gram amounts (Scheme 3.1).

Scheme 3.1. Preparation of pyrrolidine-piperidines **19**.



Surprisingly, it proved to be difficult to couple the *piperidines* **18** with tetramic acids **17** compared to the *pyrrolidines* used in syntheses of scaffolds **A**. Considerable optimization of these reaction conditions was required, as documented in Table 3.1. In the event, the highlighted conditions (entry 11: EDCI, $KHCO_3$, HOAt[80]) afforded the scaffold with two side-chains **19aa** more efficiently than any other set of conditions studied.

Table 3.1. Optimization of coupling the piperidines **18a** with the tetramic acids **17a**.



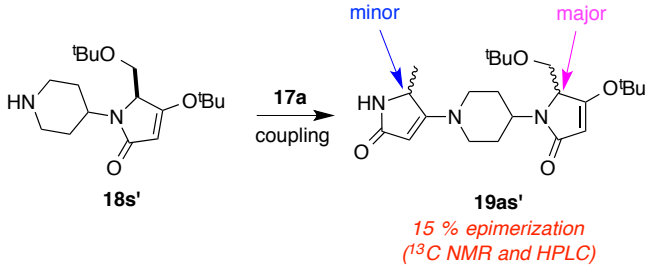
1	TBTU	TEA[a]	-	n. d.[c]	10 %
2	HBTU	TEA[a]	-	n. d.	23 %
3	PyBOP	TEA[a]	-	n. d.	complex mixture
4	TFFH	TEA[a]	-	n. d.	< 10 %
5	EDCI	TEA[a]	-	n. d.	30 %
6	EDCI	TEA[a]	HOBt	n. d.	45 %
7	EDCI	TEA[a]	HOAt	n. d.	59 %
8	EDCI	KF[b]	HOAt	n. d.	40 %
9	EDCI	Li_2CO_3[b]	HOAt	n. d.	39 %
10	EDCI	Cs_2CO_3[b]	HOAt	10 %	86 %
11	**EDCI**	**$KHCO_3$[b]**	**HOAt**	**none**	**74 %**
12	EDCI	$KHCO_3$[b]	Oxyma[81]	n. d.	complex mixture
13	EDCI	$KHCO_3$[b]	*N*-hydroxysuccinimide	n. d.	10 %

[a] 2.0 equiv. of TEA; [b] 5.0 equiv. of the base; [c] not determined.

A major difficulty arose when the optimized conditions that gave **19aa** without epimerization were applied to make **19as'** and **19fs'**; approximately 15 and 30 %

epimerization was observed by ^{13}C NMR and analytical HPLC, respectively (Scheme 3.2).

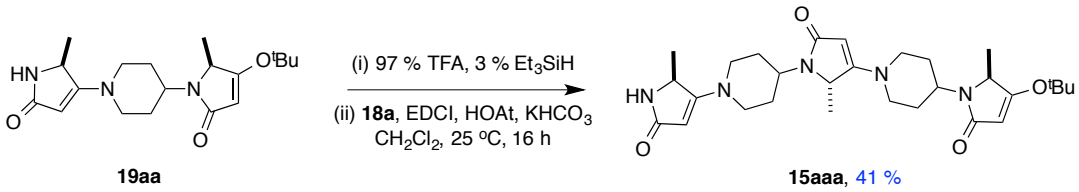
Scheme 3.2. Epimerization observed in the synthesis of **19**.



A series of experiments were undertaken to determine which of the two chiral centers was most vulnerable to epimerization. Firstly it was established that compound **18s'** is basic enough to mediate its own racemization when stored as a concentrated oil at room temperature overnight. Evidence for this assertion was from an experiment in which "aged" **18s'** was coupled with racemic tetramic acid **17a** resulting in *four* stereoisomeric products that could be separated via analytical HPLC with a chiral support (Chiralpak AD). When freshly prepared **18s'** was coupled with optically pure tetramic acid **17a**, *three* products were observed: one of these was formed in even greater amounts when the experiment was repeated with aged (partially racemized) **18s'**. The fact that a significant amount of the fourth possible stereoisomer was *not* observed in the experiments involving optically pure **17a** indicates epimerization of the **18s'** fragment dominates. A possible reason for this is inductive stabilization of the enolate from **18s'** that would not effect **17a** or the *N*-terminal part of product **19as'**.

At this stage it became clear that the targeted pyrrolidinone-piperidine oligomers **15** were too stereochemically delicate, especially if they contained side-chains that somehow promoted epimerization. Nevertheless, we were able to synthesize the three side-chain system **15aaa** without epimerization (Scheme 3.3).

Scheme 3.3. Synthesis of **15aaa**.



19aa **15aaa**, 41 %

Application of the EKOS routine indicated accessible conformations of **15aaa** contain some that overlaid on a sheet-turn-sheet motif more accurately than any of the conformers matched the other common secondary structures (3^{10}-, α-, π-helices; parallel and antiparallel β-sheets, and β-strands). Figure 3.2 shows the same overlay with conventional atoms colors on the left, and the ideal sheet-turn-sheet motif highlighted in purple on the right.

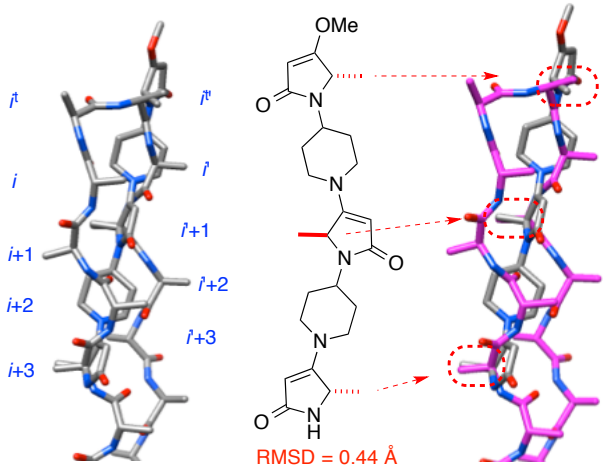


Figure 3.2. Overlay of **15aaa** on an ideal sheet-turn-sheet motif.

3.3 A General and Chemoselective Route to β-Enamino Derivatives

Our interest in β-enamino derivatives arose when pursuing the interface mimics **A, 1** and **15**. Scaffolds **A, 1** and **15** have several β-enamino amide functionalities, and overall yields in their syntheses are directly related to efficient formation of the bonds indicated. Consequently, the method for efficient formation of β-enamino derivatives under mild conditions was extensively studied in this section.[82]

β-Enamino esters are versatile synthons for many heterocycles and benzenoid compounds. There are too many applications of these building blocks to cite them all, but illustrative ones include syntheses of amino-furanones, indoles,[83-86] oxazoles,[87] pyrazoles,[88] pyridines[89] and dihydropyridines,[90] pyrimidineones,[91] pyrroles,[92-95] quinolines[96] (Figure 3.3). Existing methods to prepare β-enamino carbonyl derivatives were summarized in two papers.[97,98] These methods tend to involve acids, elevated

temperatures, and oxidizing agents, *ie* characteristics that limit the range of products that can be formed.

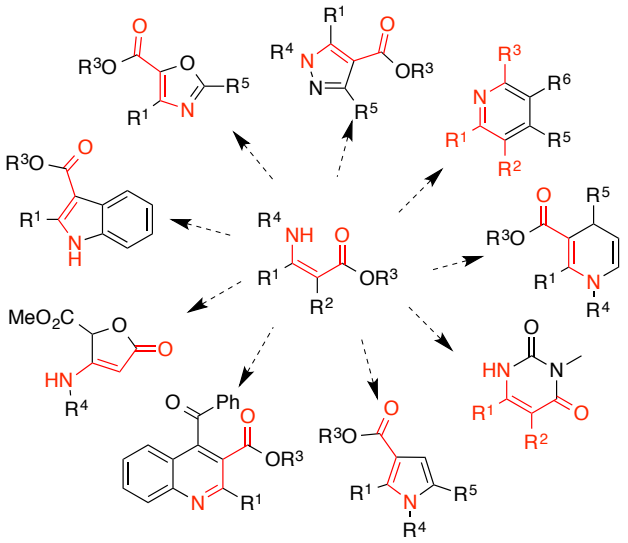
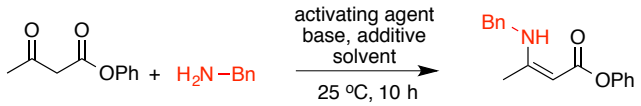


Figure 3.3. Illustrative uses of β-enamino esters.

Apparently the literature on formation of β-enamino esters from β-keto esters contains few, if any, references to conditions most commonly associated with amide bond couplings. We envisaged these types of conditions might be effective, and data from exploratory experiments to test this hypothesis are shown in Table 3.1 and Table 3.2. These pilot reactions in Table 3.2 were performed on phenolic esters to allow the opportunity for the amines to displace phenoxide, *ie* a more stringent chemoselectivity test than if less reactive esters were used. In the event, such acylation reactions did not account for significant product formation under the best conditions identified.

Table 3.2. Pilot Reactions for Formation of β-Enamino Derivatives.



entry	activating reagent	base	solvent	additive	yield (%)
1	EDC•HCl	TEA	CH_2Cl_2	HOBt	11
2	EDC•HCl	$KHCO_3$	CH_2Cl_2	-	12
3	EDC•HCl	$KHCO_3$	CH_2Cl_2	DMAP	17
4	EDC•HCl	$KHCO_3$	CH_2Cl_2	HOBt	74
5	EDC•HCl	$KHCO_3$	CH_2Cl_2	HOAt	85
6	EDC•HCl	imidazole	CH_2Cl_2	HOAt	71
7	EDC•HCl	-	CH_2Cl_2	HOAt	33
8	**EDC•HCl**	**$KHCO_3$**	**$CHCl_3$**	**HOAt**	**93**
9	EDC•HCl	$KHCO_3$	DMF	HOAt	13
10	HBTU[99]	$KHCO_3$	$CHCl_3$	-	60
11	PyBOP	$KHCO_3$	$CHCl_3$	-	80
12	**PyBOP**	**imidazole**	**$CHCl_3$**	**-**	**87**

Together, entries 1 - 5 of Table 3.2 indicate HOAt (1-hydroxy-7-azabenzotriazole)[80] is preferred as an additive over HOBt or DMAP. Entry 7, compared with all the others strongly indicates that a weak base is required. Chloroform seems to be a better solvent for this reaction type than dichloromethane or DMF (*cf* entries 8 and 12 *vs* 5 and 9). The best yield was obtained using EDC•HCl (*N*-(3-dimethylaminopropyl)-*N*′-ethylcarbodiimide hydrochloride) in entry 8, but the PyBOP ((benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate)[100] conditions in entry 12 were only marginally inferior and sometimes, depending on the substrate, have

advantages with respect to work-up simplicity. Figure 3.4 shows products and yields for other β-enamino esters formed via the method illustrated in entry 12.

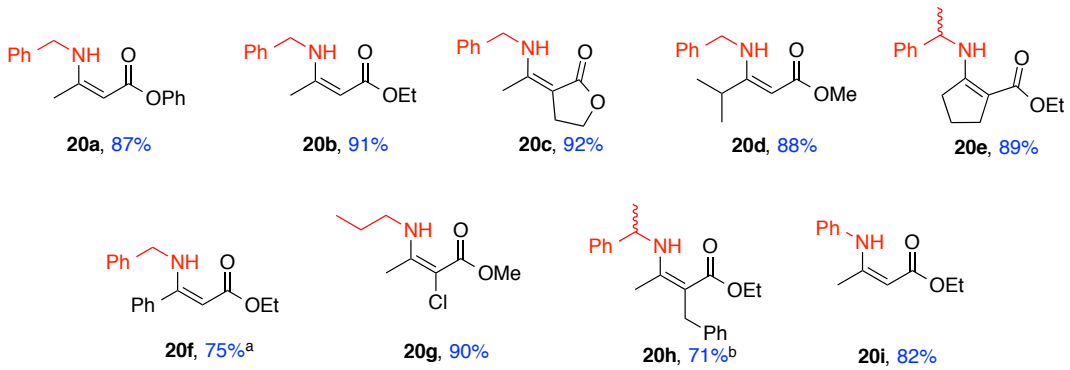


Figure 3.4. β-Enamino esters formed using the conditions shown for entry 12 in Table 3.2. Except: [a] 36 h; and, [b] 2 equiv. of amine and 24 h.

The next step in this study was to investigate β-ketothioesters as substrates. Figure 3.5 illustrates how, in the absence of additive, an amine is preferentially acylated by a β-ketothioester, rather than condensing with it. It is potentially valuable to be able to invert this chemoselectivity because that would enable preparation of β-enamino *thioesters*. Previous syntheses of β-enamino thioesters are limited to the *N*-unsubstituted forms (via multi-step procedures,[101] or by adding ammonia under buffered conditions[102]) or via condensations of *tert*-butyl β-ketothioesters in the presence of ceric ammonium nitrate (CAN) where the large *S*-alkyl group attenuates the reactivity of the thioester.[98]

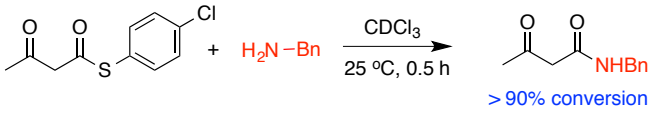


Figure 3.5. Innate reactivity of β-ketothioesters.

Inverted chemoselectivity relative to the reaction above was achieved by modifying the Table 3.2, entry 12 conditions to include the acid HOAt and hydrogen carbonate buffer. This enabled syntheses of the β-enamino thioesters shown in Figure 3.6.

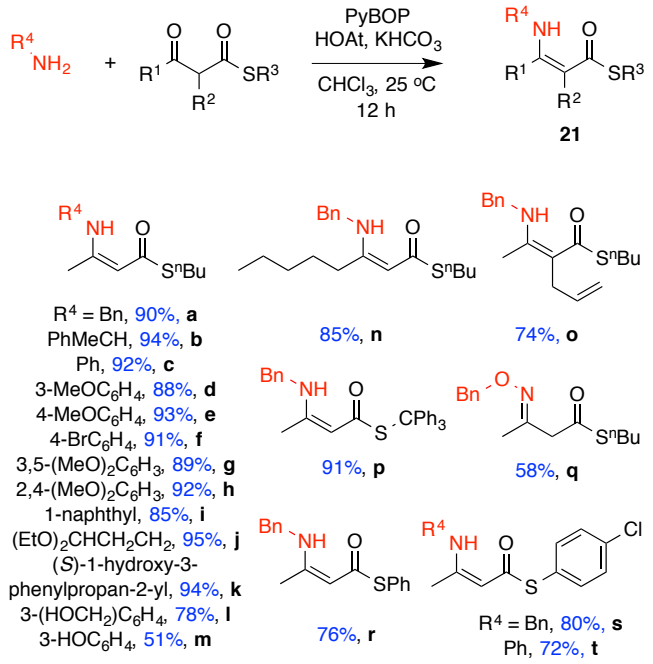


Figure 3.6. Illustrative syntheses of β-enamino *thioesters*.

Figure 3.6 demonstrates that both aliphatic and aromatic amines can be used as the nucleophile in these reactions. Formation of the β-enamino thioesters **21j** and **21p** indicates that acetal and *S*-trityl groups are tolerated, whereas they presumably would not be in the presence of acids. Products **21k**, **21l**, and **21m** demonstrate primary alcohols and phenols are compatible with the featured condensation reaction. Phenylthio ester products **21r** – **t** were formed, even though PhS^- is a better leaving group than in the other reactions.

Three Nenitzescu indole syntheses[103-105] were performed to illustrate how β-enamino thioesters can be used (Figure 3.7).[106] All three reactions did not perturb the thioester functionality, even when this involved a more reactive phenylthiolate leaving group, *ie* for indole **22b**. Formation of the regioisomers shown was confirmed via NOESY experiments.

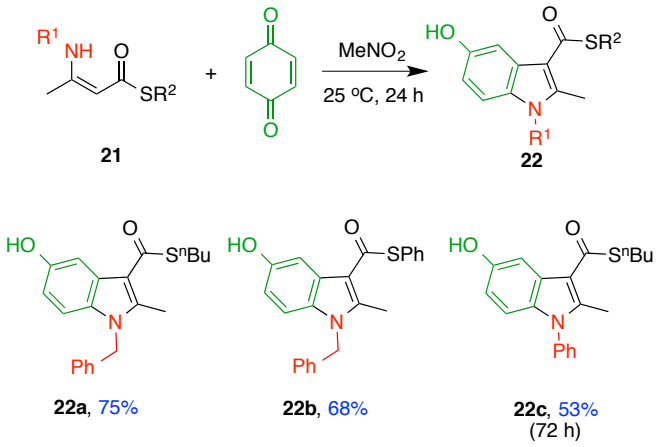


Figure 3.7. Illustrative Nenitzescu indole syntheses.

An isolable intermediate azabenzotriazole ether was formed when 1,3-cyclohexandione was reacted under the conditions used under typical coupling conditions (Figure 3.8). It seems probable that this intermediate is more stable than the corresponding one in the thioester reactions. If similar intermediates are formed throughout, this reaction indicates a two-step conjugate addition-elimination process is operative. This would also explain why HOAt is superior to HOBt in these transformations: because the aza-derivative can "lever" proton transfer in the amine addition step as indicated in transition state in Figure 3.8.

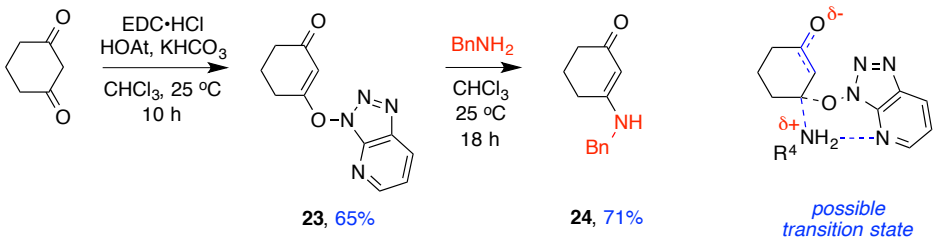


Figure 3.8. Intermediate formation in HOAt-mediated reactions.

Finally, Figure 3.9 illustrates how the conditions developed for formation of β-enamino amides were applied in reactions typical of those used to make the interface mimics.[107] Good yields were obtained, and the near-neutral conditions indicate there is likely to be a broad substrate scope.

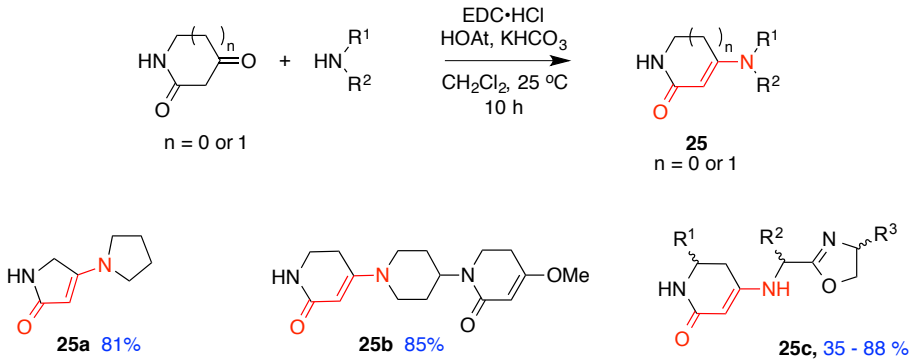


Figure 3.9. Application of the featured conditions for syntheses of interface mimic fragments.

In summary, β-enamino derivatives can be formed via carefully buffered conditions featuring activation agents like EDC•HCl and PyBOP. These conditions are sufficiently mild to facilitate synthesis of a wide range of examples, including relatively delicate synthons like β-enamino thioesters.

3.4 Conclusions

In this chapter, a minimalist mimic based on pyrrolidinone-piperidine scaffold was designed and synthetic route to the mimic was explored. The mimic can be synthesized but the chiral centers are too vulnerable to epimerization. Conformational studies indicated that the scaffold could mimic a sheet-turn-sheet motif better than the other common secondary structures. The synthetic methodology developed to prepare oligo-pyrrolidinone-piperidine was thoroughly explored and extended to the synthesis of other β-enamino derivatives. This method enabled the chemoselective synthesis of a

series of β-enamino thioesters, which cannot be prepared efficiently with other currently existing methods.

CHAPTER IV

OLIGO-PIPERIDINE-PIPERIDINONE PROBES THAT PERTURB A PROTEIN-PROTEIN INTERFACE IN ANTITHROMBIN*

4.1 Introduction

Serpins are proteins that, in their monomeric forms, function as naturally occurring serine protease inhibitors. Monomeric serpins are metastable,[108] and they revert to thermodynamically more favorable (*ca* 32 kcal/mol)[109] dimeric, then oligomeric, forms. Oligomeric serpins assemble into fibrils that are associated with a series of diseases known as "serpinopathies"[110] encompassing conditions as diverse as some neurological conditions and liver sclerosis.

α-Antithrombin is a serpin that inhibits thrombin. It played an important role in the development of understanding how serpins may form fibrils because Huntington *et al* were able to crystallize a physiologically relevant dimer that now serves as a model for the formation of oligomers (the oligomers have not yet been structurally characterized on a molecular level, presumably because they are non-homogenous involving a range of molecular masses). Figure 1 illustrates how the red and purple sheet regions of the α-antithrombin monomer (PDB: 2ANT) reorganize to form a purple cleft to accept an interface hairpin from the partner, and donate a red hairpin to it in the dimer (2ZNH).[109,111] Consequently, one of the several models for serpin oligomer

*Reproduced in part from "Small Molecule Probes That Perturb a Protein-protein Interface in Antithrombin", Dongyue Xin, Andreas Holzenburg and Kevin Burgess, *Chem. Sci.* **2014**, *5*, 4914-4921. Copyright 2014 The Royal Society of Chemistry.

formation[109,111-118] is via domain swapping to form a dimer, then repetition of this process to form oligomers.[117,118] Thus serpin-dimer formation is potentially critical in serpinopathies, and has been described as "infectious".[112,119,120]

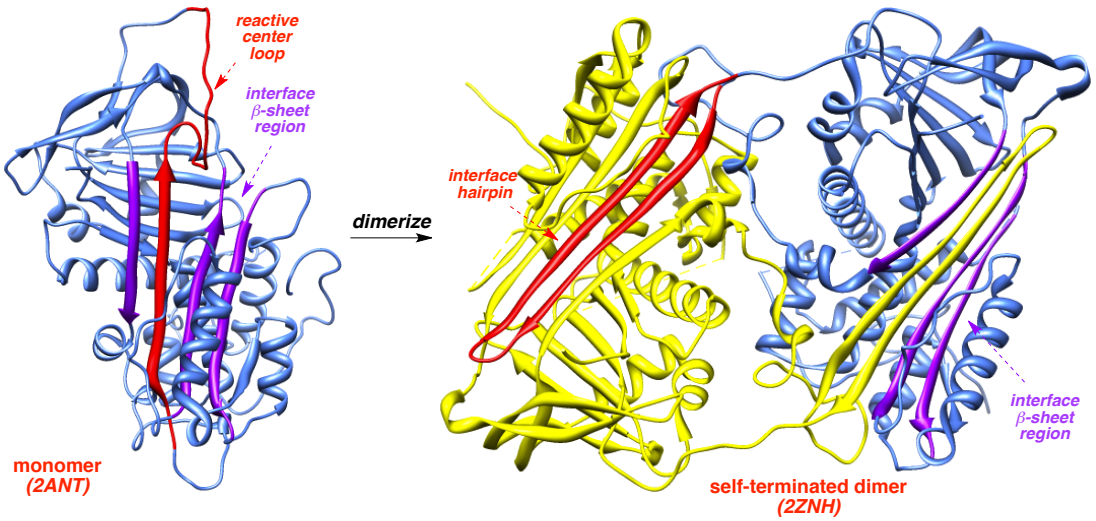


Figure 4.1. Structures of the α-antithrombin monomer and the self-terminated dimer.

Scaffolds **1**, as featured in Chapter II, are a good example of the kind of chemotype that can be processed using the EKO approach. Molecules in this series contain three β-amino acid fragments (blue) that can be made from the corresponding α-amino acid chirons, and have only two significant degrees of freedom (red arrows). (Figure 4.2)

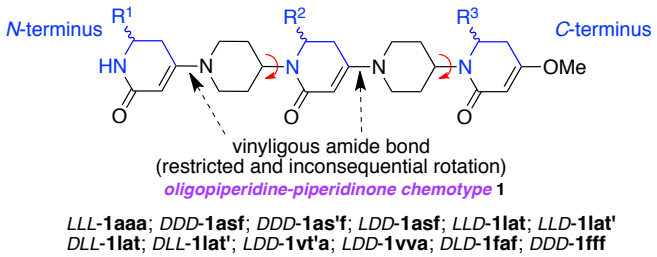


Figure 4.2. Oligo-piperidine-piperidinone probes studied.

Synthetic protocols were described for compounds **1** in Chapter II.[65] EKOS was used to compare the simulated conformations of all isomers of **1aaa** with ideal secondary structures; this indicated that stereomers of chemotypes **1** can adopt conformations that resemble all common secondary structures, with a bias towards extended sheet-turn-sheet and β-sheet conformations more than helical ones.[65]

For the work described in this chapter, EKO identified several PPI interface regions that correspond to side-chain orientations found in synthetically accessible conformations of **1**. Consequently, we set out to ascertain if the compounds **1** implicated in this EKO analysis would in fact perturb PPIs that influence α-antithrombin oligomerization. These experiments were not performed as part of a drug discovery effort, but rather to test if EKO could be used to identify compounds that perturb PPIs in this particular test case. Specifically, the goal was to explore if the structures implicated by EKO would perturb PPIs in antithrombin and impact (positively or negatively) the oligomerization process.[121]

4.2 Perturbation of α-Antithrombin Oligomerization by Interface Mimics

Results from EKO imply scaffold **1** might present side-chains in several ways that resemble the α-antithrombin dimer PPI interface; the four that were experimentally tested in this work are listed in Table 4.1, and three representative overlays are shown in Figure 4.3. One (RMSD 0.26 Å) involves three residues on a single strand of the β-hairpin in an *i*, *i* + *2*, *i* + *5* relationship (*i.e.* ^{384}Ala, ^{386}Thr, and ^{389}Val; Figure 4.3a and entry 1 in Table 4.1). Conversely, the scaffold spanned the *two* strands of the β-hairpin in other overlays, matching ^{373}Leu and ^{371}Ala on one strand and ^{386}Thr on the other in Figure 4.3b, and in Figure 4.3c ^{383}Ala/^{385}Ser on one strand and ^{368}Phe on the other. Another issue surrounding data mining within the EKO strategy relates to the *polarity* of the overlays. Scaffolds like **1**, being formed from amino acids, have recognizable *N*- and *C*-termini, and these can be overlaid parallel or antiparallel to the protein. In Figure 4.3a and b the overlay is *antiparallel* so the mimics that should be prepared are LDD-**1vta** and LLD-**1lat**. Figure 4.3c however, shows mimic conformers overlaid on the strands in a *parallel* fashion hence the target is DDD-**1asf** and not DDD-**1fsa**. Like entry 2, the fourth overlay referred to in Table 4.1 (entry 4, but not shown in Figure 4.3) also corresponds to the **1lat** sequence, but the stereochemistry and polarity is different.

Table 4.1. Conformational matches identified by EKO on interface regions in the α-antithrombin dimer structure.

entry	isomer	RMSD(Å)	residues	polarity
1	LDD	0.26	V389-T386-A384	antiparallel
2	LLD	0.33	L373-A371-T386	antiparallel
3	DDD	0.35	A383-S385-F368	parallel
4	DLL	0.37	L373-A384-T386	parallel

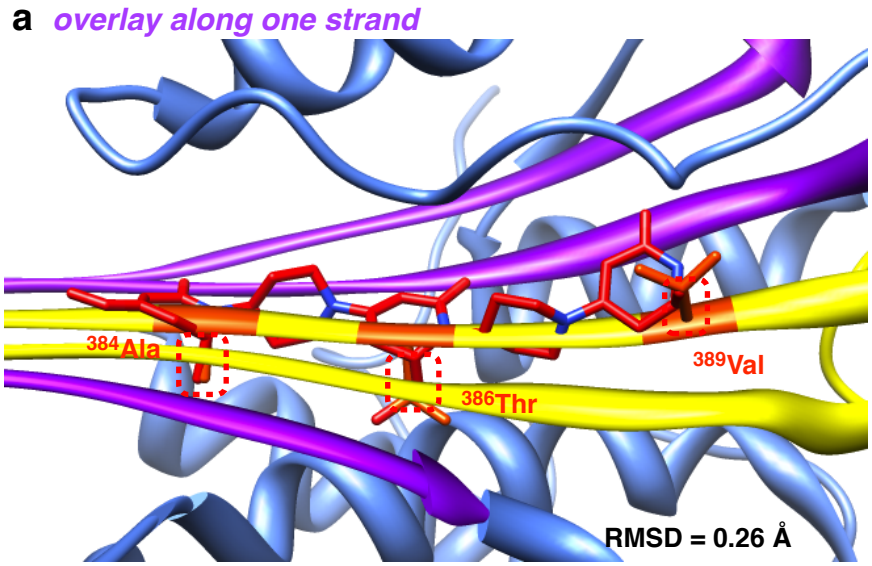


Figure 4.3. Conformers of scaffolds **1** can overlay side-chains on the β-hairpin structure in the α-antithrombin dimer. **a** on one strand; or, **b** and **c** spanning across two strands.

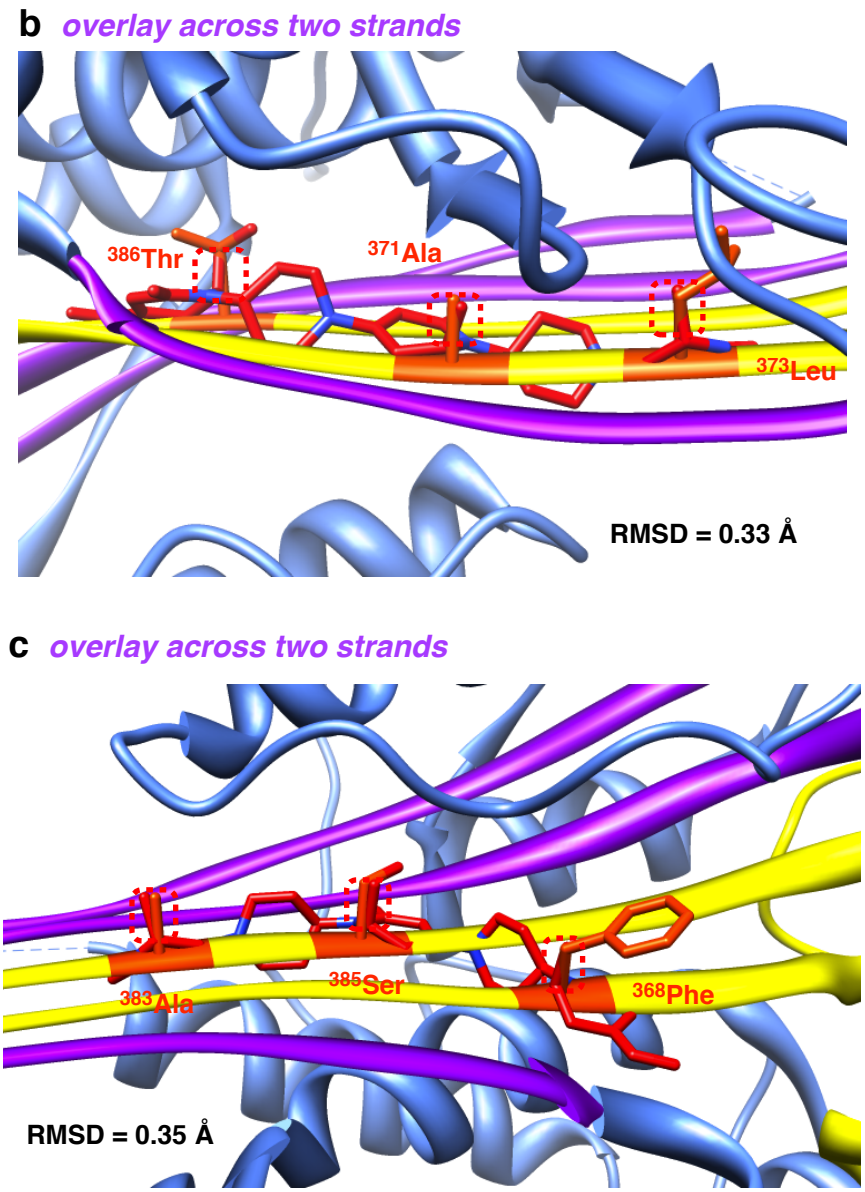


Figure 4.3. Continued.

The compounds specified in Table 4.1 were prepared to test the hypothesis that EKO can implicate small molecules to disrupt PPIs in the α-antithrombin dimer,[109] but with one exception. Entry 1 of Table 4.1 calls for synthesis of DDL-**1vta**. However, it is considerably easier to make DDL-**1vva**, with Val replacing the Thr, because of issues with side-chain hydroxyl protection. Consequently, we elected to prepare DDL-**1vva** in place of DDL-**1vta** on the basis that the shape of the Val side-chain ($CHMe_2$) is similar to that of Thr (CH(OH)Me).

The next step in the process was to devise assays to test perturbations of PPIs in α-antithrombin dimer. α-Antithrombin oligomerizes more readily at temperatures above ambient, and at low pH values. After some experimentation, it was shown that such oligomerizations occur at convenient rates for our purposes above 50 °C at pH 7.4. When α-antithrombin was incubated at 50 °C for 8 h at pH 7.4, it showed little sign of oligomerization as monitored via native PAGE (*i.e.* non-denaturing gel; Figure 4.4a). Experiments were conducted to determine the effects of control or target compounds being present during incubation (Figure 4.4b, where target compounds are shown in red and controls or "partial controls" are shown in blue). Thus, lane 2 of Figure 4.4b shows that the scaffold with three methyl side-chains (LLL-**1aaa**) gave a small amount of oligomerization; this was unsurprising because all the EKO-implicated targets in Table 4.1 have a methyl side-chain and share the same scaffold, so LLL-**1aaa** is too similar to be a true negative control. However, tests with indigo, a compound that is known to bind many proteins non-specifically,[122] showed no evidence that it catalyzed oligomerization (see Figure 4.4d).

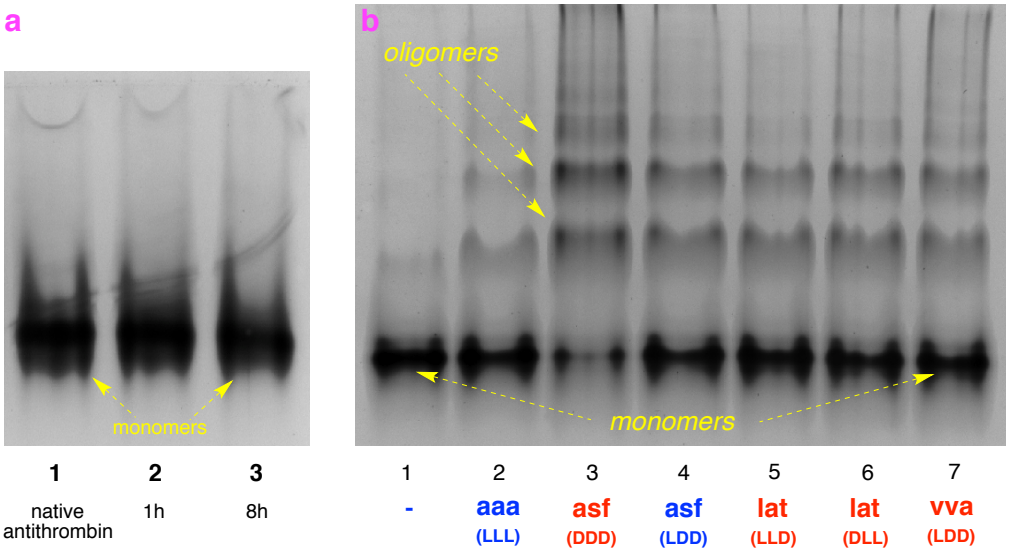


Figure 4.4. Native gel electrophoresis study of α-antithrombin oligomerization. Throughout, controls are delineated in blue, and assays of target compounds are shown in red. After 1 h at 50 °C in pH 7.4 at 0.25 mg/mL concentration, α-antithrombin shows the following. **a** α-Antithrombin was incubated at 50 °C for 0 h (lane1), 1 h (lane2), and 8 h (lane3). **b** (Lane): 1 little or no oligomerization on its own; and, 2 only slightly more in the presence of the scaffold LLL-**1aaa** (only methyl side-chains). However, lane 3 shows a target compound implicated in EKO, DDD-**1asf** catalyzes formation of significantly more oligomers, particularly the ones having lower molecular mass. Lane 4 shows oligomer formation is slightly suppressed for LDD-**1asf** relative to DDD-**1asf** (corresponding to one inverted stereocenter). Lanes 5 – 7 show the indicated target compounds implicated by EKO also induce oligomerization, though less than DDD-**1asf** under the same conditions. **c** A similar comparison for some *benzyl-protected* target compounds (red) and protected or otherwise hydrophobic controls (blue). 200 fold of indicated compounds were used in all cases. **d** α-Antithrombin (0.25 mg / mL) was incubated with 200 fold compound LLL-**1aaa** (negative control in this series, lane 1), 200 fold DDD-**1as'f** (lane 2), 200 fold indigo (control for non-specific hydrophobic interactions, lane 3), 0.025 % Triton-X 100 (control for surfactant, lane 4), 200 fold sodium dodecyl sulfate (SDS, control for amphiphilic compounds, lane 5) at 50 °C for 1 h.

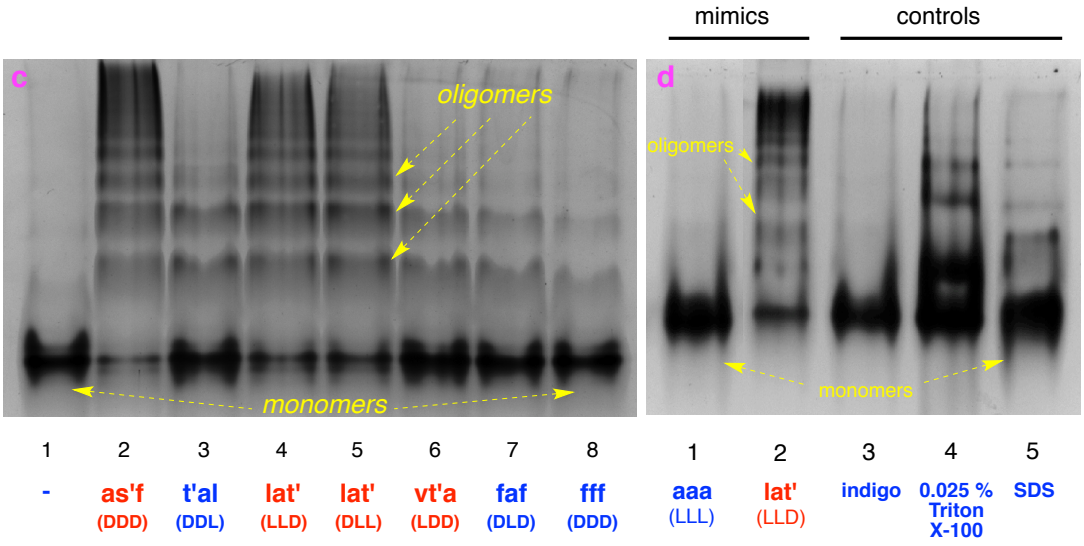


Figure 4.4. Continued.

Lanes **3**, **5** – **7** in Figure 4.4b showed that the four compounds targeted (DDD-**1asf**, LLD-**1lat**, DLL-**1lat** and LDD-**1vva**, in red) imparted significantly more oligomerization than LLL-**1aaa**. Lane 4 shows data for another interesting "partial control", LDD-**1asf**. EKO did *not* indicate that conformers of this compound would match the featured PPI interface, but it might be expected to do so to some degree since it is an epimer of ones that EKO did implicate (DDD-**1asf**, lane **3**). In the event, both LDD-**1asf** and DDD-**1asf** do mediate oligomerization, but the target compound DDD-**1asf** did it more effectively than the partial control LDD-**1asf**. Finally, the gel in Figure 4.4b shows that all the conditions give oligomers bias towards the lower molecular mass range.

Figure 4.4c shows a gel that compares *benzyl-protected* forms of the targeted compounds with some similarly hydrophobic derivatives of the same scaffold **1**, but which have side-chains and/or stereochemistries that were not implicated by EKO. Lane

2 shows DDD-**1as'f**, the benzyl-protected derivative of the most active compound in Figure 3a, also catalyzed the oligomerization process. The "partial control" DDL-**1t'al** in lane 3 has two amino acids interchanged relative to the protected, targeted compound LLD-**1lat'**, and a different stereochemistry; the protected targeted compound accelerated the oligomerization most. Moreover, another protected, targeted compound having the same sequence of amino acids, DLL-**1lat'** in lane 5, also gave more acceleration than the control with two side-chains switched, DDL-**1t'al**. Finally, LDD-**1vt'a** is the only protected, target compound that did not markedly accelerate the oligomerization process. The hydrophobic compounds DLD-**1faf** and DDD-**1fff,** having the same scaffold but no side-chain correspondence, also did not promote oligomer formation significantly.

All the PAGE data presented above are based on qualitative silver-stains. Kinetic data from the band intensities were determined to quantitate and compare effects of the compounds.[123,124] Figure 4.5a shows an illustrative data correlation of rate constants for oligomerization of α-antithrombin with the concentration of mimic DDD-**1as'f**. In Figure 4.5b, rate constants for the same oligomerization but in the presence of five featured mimics are compared; DDD-**1as'f** was the best catalyst. Kinetics was not performed on all the *unprotected* mimics shown in Table 4.1 because it was apparent that the protected forms tend to be better catalysts.

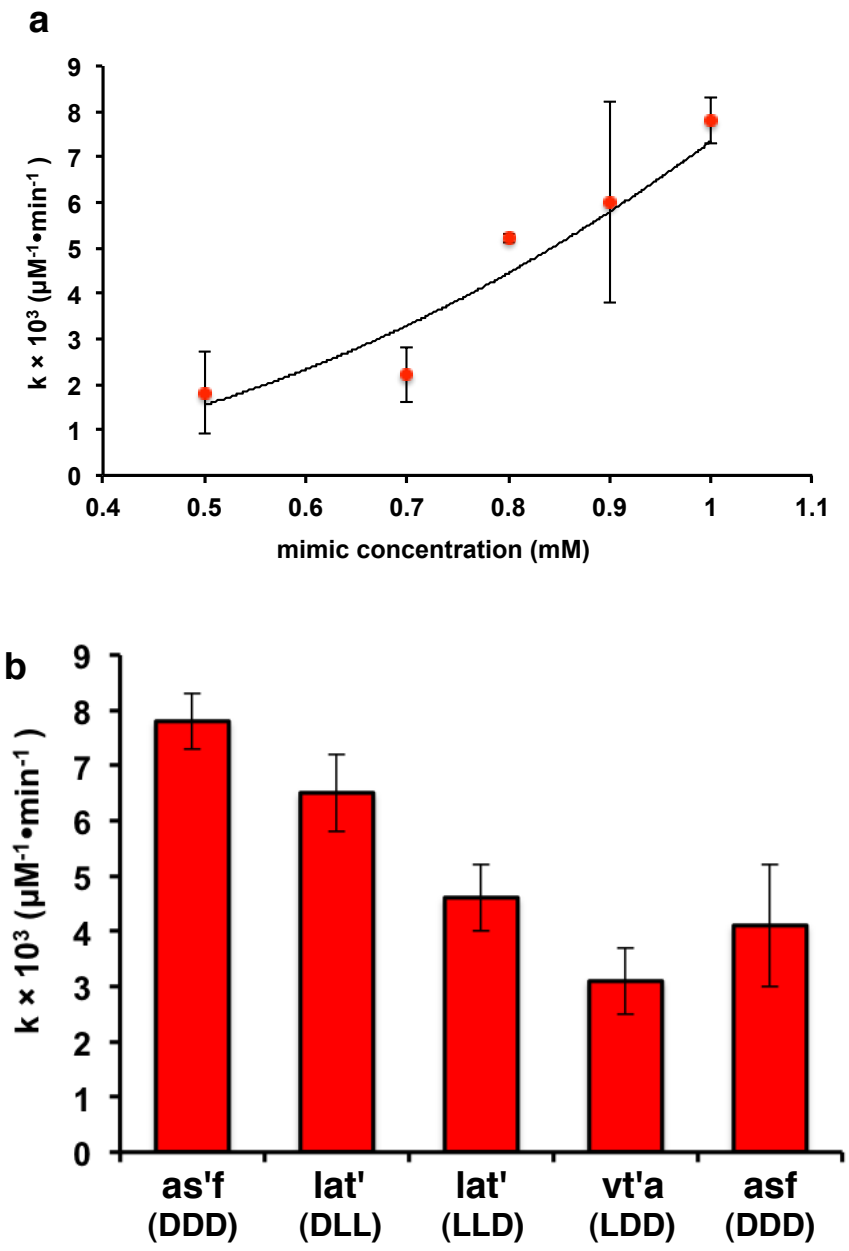


Figure 4.5. Kinetics of α-antithrombin oligomerization. Induced by **a** DDD-**1as'f**; and, **b** benzyl-protected forms of the four featured mimics, and the best unprotected compound, DDD-**1asf** (all at 200x the protein concentration).

Another way to assay these compounds was to test antithrombin samples after oligomerization for their residual activity as thrombin inhibitors. Thus, inhibition of thrombin by α-antithrombin (optionally, after putative oligomerization mediated by the small molecule probes **1**) was used to assay the featured compounds. Thus, the probes were incubated with α-antithrombin (time variable), and aliquots of this solution were used to inhibit thrombin-mediated hydrolysis of a peptide containing 4-nitroaniline (Spectrozyme® TH, from American Diagnostica); this type of assay is a standard test for thrombin activity (performed in the presence of heparin).[125,126] In the event, thrombin was most active when inhibited by samples of α-antithrombin that had been incubated with DDD-**1as'f** (Figure 4.6a). This observation is consistent with the assertion made above, that DDD-**1as'f** promotes α-antithrombin *deactivation* by oligomerization well. When α-antithrombin deactivation in this assay was compared for DDD-**1as'f** and DDD-**1asf**, the benzyl-protected compound induced significantly more deactivation.

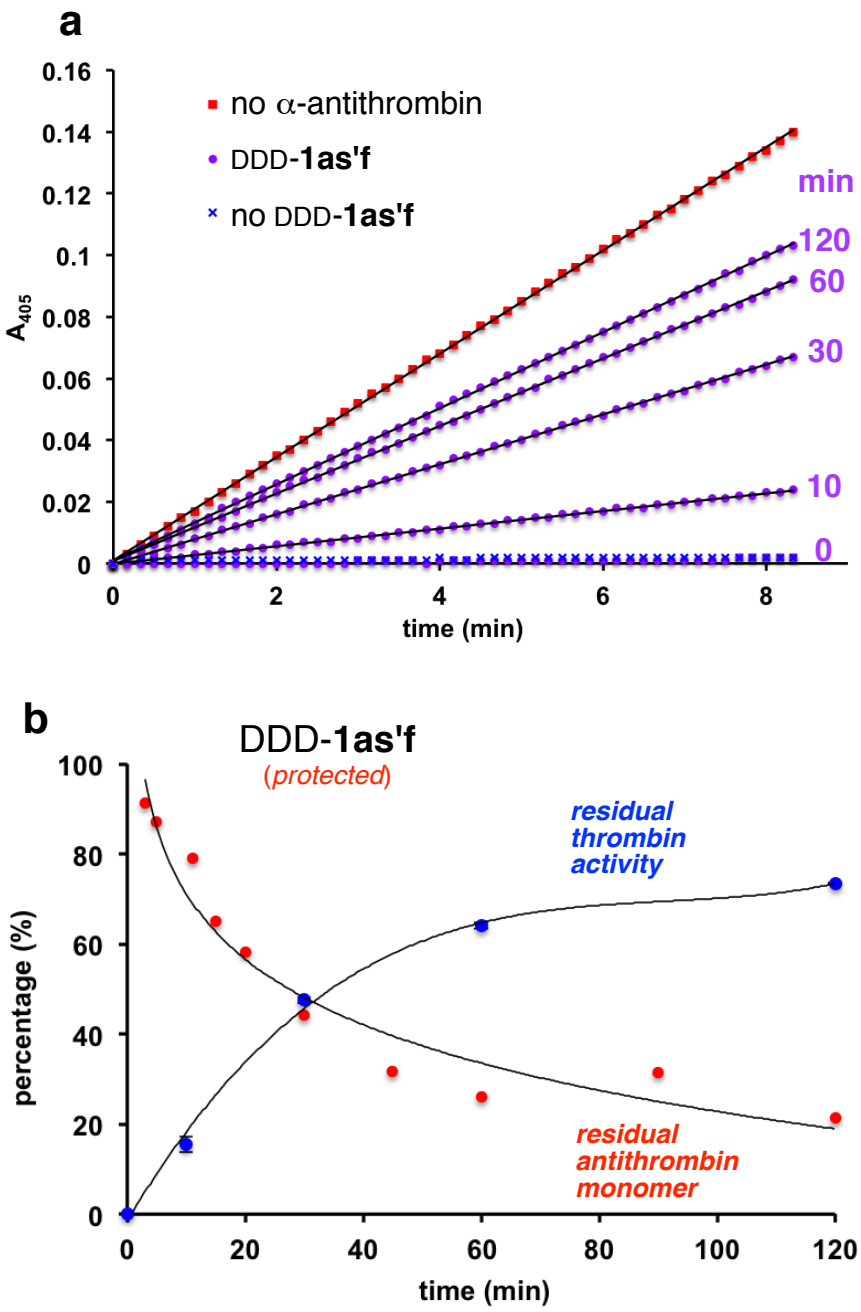


Figure 4.6. Activity of antithrombin as thrombin inhibitors revealed by enzyme kinetics studies. **a** Absorption of 4-nitroaniline formed by thrombin, and inhibited by α-antithrombin that was previously incubated with DDD-**1as'f** for the times shown. Correlation of residual thrombin activities with α-antithrombin monomer concentrations are shown in: **b** for DDD-**1as'f**; and, **c** for DDD-**1asf**.

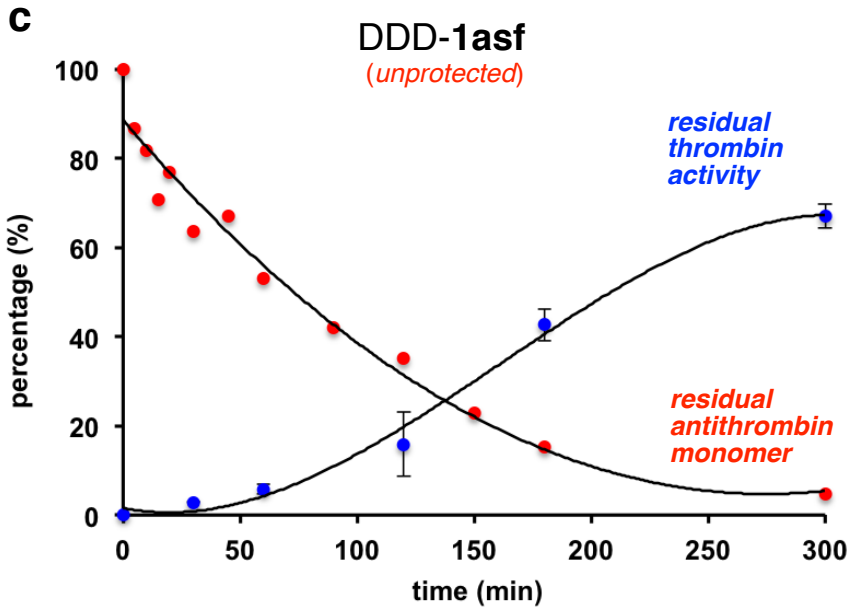


Figure 4.6. Continued.

A subtle difference in the data depicted in Figure 4.6 is that 50 % residual α-antithrombin monomer in 4.6b (for DDD-**1as'f**) corresponds to *high* thrombin activity, but in 4.6c (for DDD-**1asf**) 50 % residual α-antithrombin monomer corresponds to *low* thrombin activity; that observation appears to be counterintuitive. However, recall from Figure 4.4 the protected probe DDD-**1as'f** favors formation of *longer* α-antithrombin oligomers than the deprotected form, DDD-**1asf**. In general, formation of oligomers involves burying the free "reactive center loop" that complexes to thrombin to give inhibition in a protein cavity. α-Antithrombin oligomers have a free reactive center loop at one end, which may *still* inactivate thrombin even though the protein is oligomerized. In short oligomers, the proportion of free reactive center loops available to inactivate

thrombin is *higher* than in long oligomers. Thus data in Figures 4.6b and c can be reconciled by accounting for the proportion of uncomplexed reactive center loops in the oligomeric products.

Alterations of protein tertiary and quaternary structures are often followed using circular dichroism (CD) and isothermal calorimetry (ITC). In this study, addition of the compounds to α-antithrombin under conditions that were expected to cause oligomerization resulted in changed CD spectra, as anticipated.[120] However, the data was hard to interpret beyond this crude observation because transformation of an α-antithrombin tertiary structure to another similar one in the dimeric and oligomeric forms involves only small changes in elipticity. When α-antithrombin (at 20 μM) was mixed with the optimal probe, DDD-**1as'f**, under conditions up to a 1:1 ratio in an ITC bomb, the enthalpy change could not be detected. This also is unsurprising because the binding of the probe to the protein may not liberate sufficient enthalpy to detect in ITC under these conditions. Overall, these experiments are hard because they are not simple binding events, but instead the molecule acts as a catalyst to induce different oligomerization states.

Whereas CD and ITC are apparently inappropriate to follow induced α-antithrombin oligomerization, electron microscopy enabled direct visualization of the process. Negatively stained untreated antithrombin molecules revealed a range of projections and a representative area is depicted in Figure 4.7 (inset). In the most frequent projection, the molecules assume a roughly elliptical shape. Molecular mass can be determined from the dimensions of these particles, if a 3D shape can be inferred.

Assuming a prolate ellipsoid as the overall 3D shape and a partial specific volume (υ) of 0.74 ml/g[127,128] the molecular mass (m) can be calculated according to the formula m (Da) = volume of the protein (ml) x Avogadro's number x 1/υ.[128,129] Using this formula together with the long and short half axes determined as 3.1 ± 0.3 nm and 2.3 ± 0.3 nm, respectively, one arrives at a molecular mass of 56 kDa. This finding suggests that the imaged untreated particle constitute a monomeric antithrombin population as the published molecular mass for antithrombin is 58 kDa.[109] The dramatic change in appearance upon treating antithrombin with the mimic DDD-**1as'f** is readily apparent in Figure 4.7 which is characterized by a large population of higher order oligomer. α-Antithrombin oligomerized by treatment with the optimal probe (DDD-**1as'f** at 50 °C) appears as small fibrils as shown in Figure 4.7.

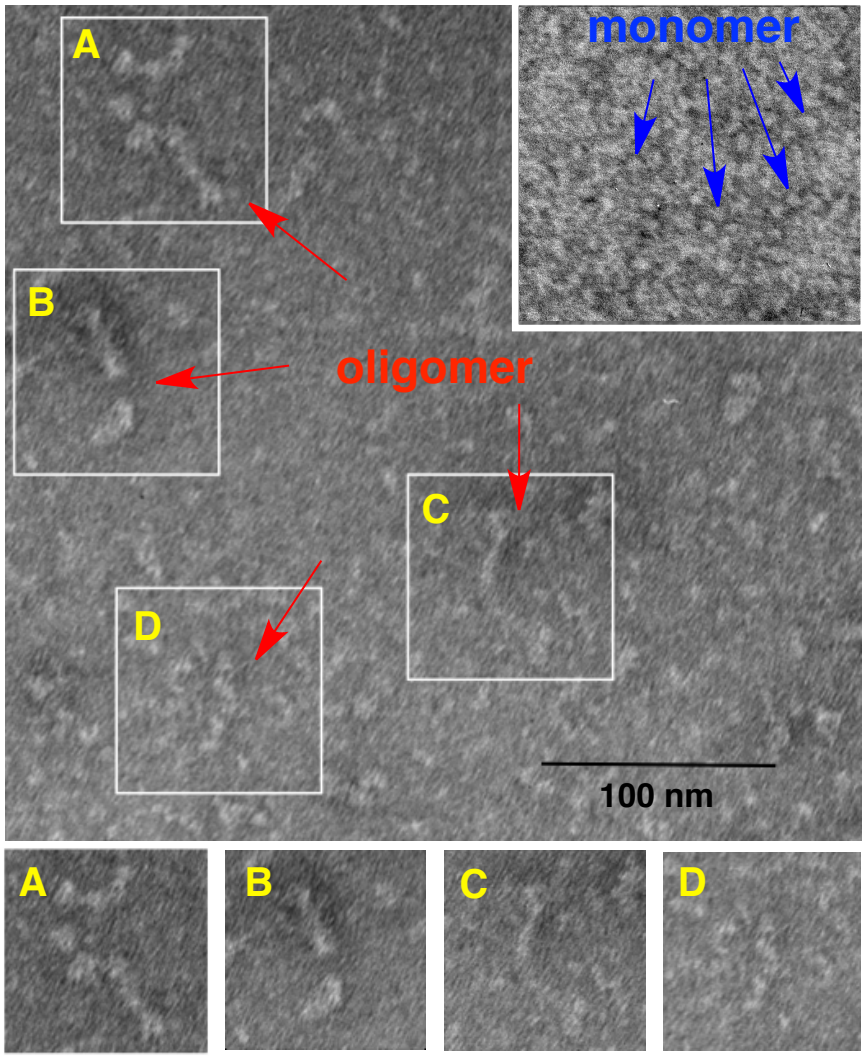


Figure 4.7. Electron micrographs of negatively stained α-antithrombin. Prior to any polymerization conditions (inset); and incipient α-antithrombin fibrils induced by the action of DDD-**1as'f** at 50 °C for 1 h. Expansions A – D highlight regions of interest where the oligomers can be directly observed.

Finally, the featured compounds were designed for perturbation of the dimer intermediate in Figure 4.1, but the evidence outlined above would also be consistent with the compounds acting via perturbation of the monomer. This motivated us to consider how the monomer might fare in the EKO analysis, but there was one critical problem.

The algorithm that EKO uses is specifically designed for PPIs involving more than one chain, so it will not "pick up" hits on the antithrombin *monomer* interface. Consequently, to answer this question, we tricked the algorithm by breaking the monomer PDB so that it appears to be a PPI involving two chains. When the EKO process was applied to that *pseudo*-dimer, hits were found on the pivotal β-sheet region corresponding to in the chain swapped dimer.

4.3 Conclusions

Application of the EKO strategy to chemotypes **1** led us to hypothesize that selected probes in this series would perturb PPIs in α-antithrombin. In the event, four target compounds, and four benzyl-protected precursors, were tested, and all catalyzed oligomerization of α-antithrombin. This conclusion is supported by observation of oligomers on gels, by monitoring residual α-antithrombin inhibition activity in thrombin assays, and via direct observation using EM. These findings are consistent with the original hypothesis that the targeted probes perturb PPIs in α-antithrombin.

An interesting outcome of this study is that benzyl-protected forms of the target compounds proved to be better oligomerization catalysts than the non-protected compounds. This does not appear to be a non-specific hydrophobic effect because true negative controls (*eg* indigo), and closely related compounds with conformational states that were not marked by EKO as being appropriate, did not catalyze the oligomerization to the same extent. Figure 4.8 overlays based on EKO for two benzyl-protected target compounds (one protein removed and only protein-binding partner remains). To explain

why DDD-**1as'f** is more active than the deprotected form we propose that the side-chain benzyl (shown as a red sphere in 4.8a) precludes the dimer protein conformation via a clash with the strand-helix motif shown on the right of that graphic. In support of this, literature observations indicate disruption of that region may promote unfolding in the oligomerization process (based on α-antithrombin mutagenesis experiments).[130] Figure 4.8b suggests the binding of LLD-**1lat'** may be favored by placing the side-chain benzyl in a hydrophobic pocket.

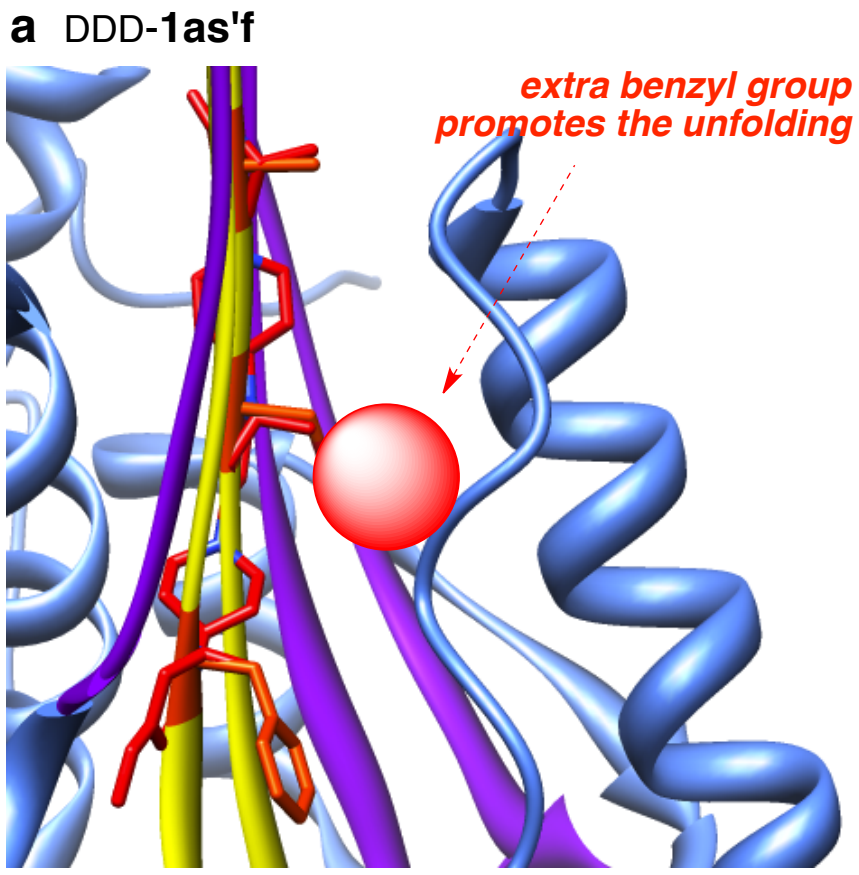


Figure 4.8. Proposed binding modes based on EKO analyses. **a** DDD-**1as'f**; and, **b** LLD-**1lat'**.

b LLD-**1lat'**

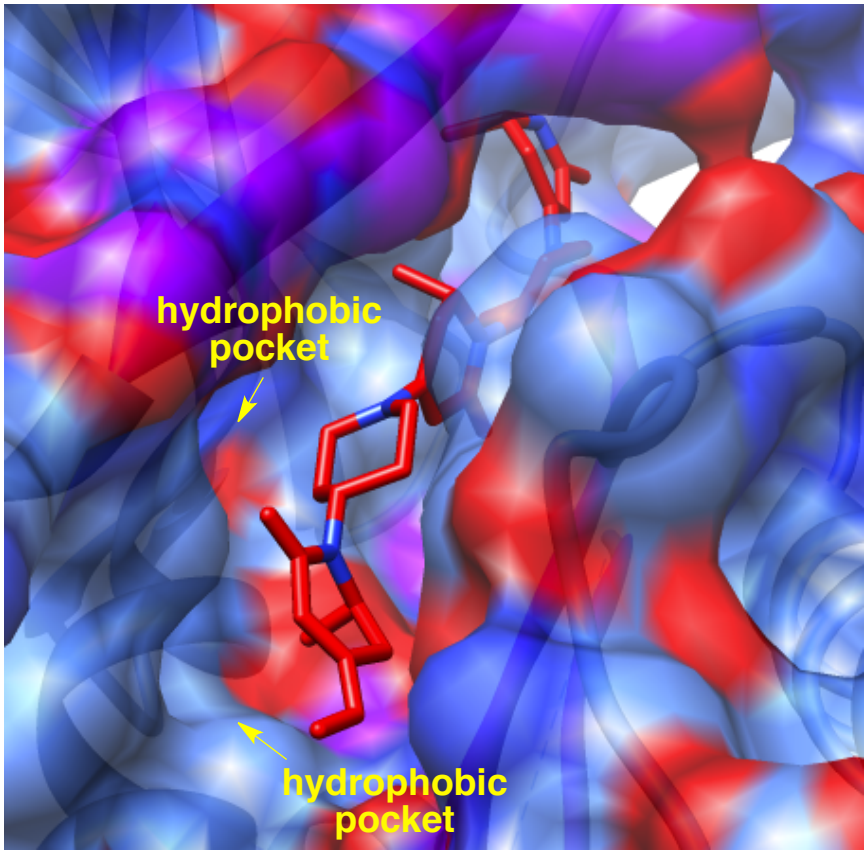


Figure 4.8. Continued.

Sheet and strand mimics have been reported in the literature,[60,131-135] but, to the best of our knowledge, no small molecules have been reported to perturb the oligomerization of α-antithrombin. One report described small molecules that perturb oligomerization of *another* serpin, a mutant of α_1-antitrypsin or "Z α_1-antitrypsin",[117] but this was discovered via virtual screening of molecules to fit in a cavity; molecules docking with that site appear to cause an allosteric interaction.[136] However,

corresponding cavity is not present at the α-antithrombin PPI interface,[136] so exactly the same type of allosteric interaction is impossible for this protein. Unfortunately, elucidation of the protein regions where the small molecules impact would be hard due to the transient nature of the binding, and the heterogeneous mixture of proteins formed in α-antithrombin oligomerization; consequently, catalysis of oligomerization via docking of the probes with an allosteric site not at the PPI interface cannot be ruled out at this stage. However, perturbation of serpin oligomerization via compounds that act at the PPI interface is possible since peptides based on this region have that effect,[137] including ones that impact Z α_1-antitrypsin[138,139] or α-antithrombin.[140,141]

The EKO strategy applied in this study led to compounds that promoted the oligomerization of α-antithrombin. EKO has no provision to determine what the biochemical effects of disrupting a particular PPI may be; in the event the perturbation here promoted α-antithrombin oligomerization.

CHAPTER V

CYCLIC PEPTIDES AS TURN MIMICS

5.1 Introduction

Cyclization of linear peptides increases their proteolytic stabilities and rigidities. In ideal cases these structures will adopt only *one* preferred conformation; if that occurs, less entropy will be surrendered on interaction with biomolecular receptors, increasing the free energies for the interactions. Observation of a strongly preferred conformation in solution also makes it probable that the molecule will bind to the receptor in a similar conformation, compared to other situations in which the compound exists in several solution conformations. Moreover, exclusion of competing conformational states reduces possibilities for off-target binding. In the context of our studies on EKO, hits from conformationally rigid hits are potentially more interesting than hits from more flexible compounds because the entropy of binding is more favorable. Moreover, cyclic peptides are routinely made via solid phase methods, and, after laboring to prepare oligopiperidine-pyrrolidionone chemotypes, we realized how critical convenience synthetic accessibility is.

Inconveniently, cyclic peptides composed of the 20 genetically encoded amino acids miss a "sweet spot" ring size where conformational homogeneity is attained without compromising ease of syntheses. By this we mean, cyclic *tri*-[142-145] **A** and *tetra*-peptides **B**[146-152] are notoriously difficult to prepare because they are constrained in 9- and 12-membered ring conformations. Analogs of cyclic tetrapeptides, like the 12-

membered ring system **C**, may be more easily prepared but another problem arises: *cis/trans* amide bond equilibria introduces conformational heterogeneity.[153] Cyclic *penta*peptides,[154-156] are easier to make than cyclic tri- or tetrapeptides[157] because their 15-membered rings are less strained, but they tend to equilibrate between conformers **D**$^{1-5}$ containing β- and γ-turns. Certain states in the **D**$^{1-5}$ equilibrium can be favored if one of the amino acids has a D-configuration, especially D-Pro,[158] but most cyclic pentapeptides and higher homologs overall do not tend to be rigid unless further constrained.[159,160]

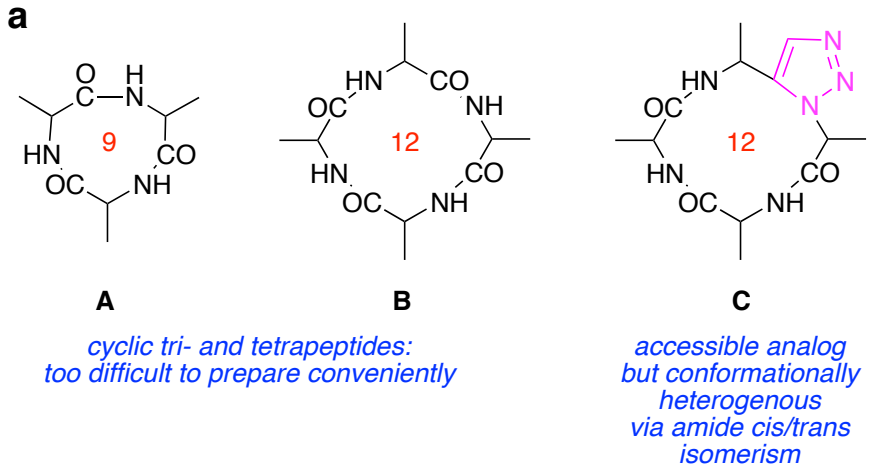


Figure 5.1. Representative previously reported cyclic peptide systems. **a** 12-Membered ring peptidic systems (*eg* cyclic tetrapeptides) are highly constrained and their conformations may be complicated by *cis-trans* isomerism. **b** Cyclic pentapeptides tend to equilibrate between similar states containing both γ and β-turns, *ie* they tend to be conformationally heterogeneous. **c** 13-Membered rings may give only one preferred conformer (*eg* **E** - **G**) but this issue has not previously been studied for systems containing the *Anth* residue (*eg* for **H** and **I**).

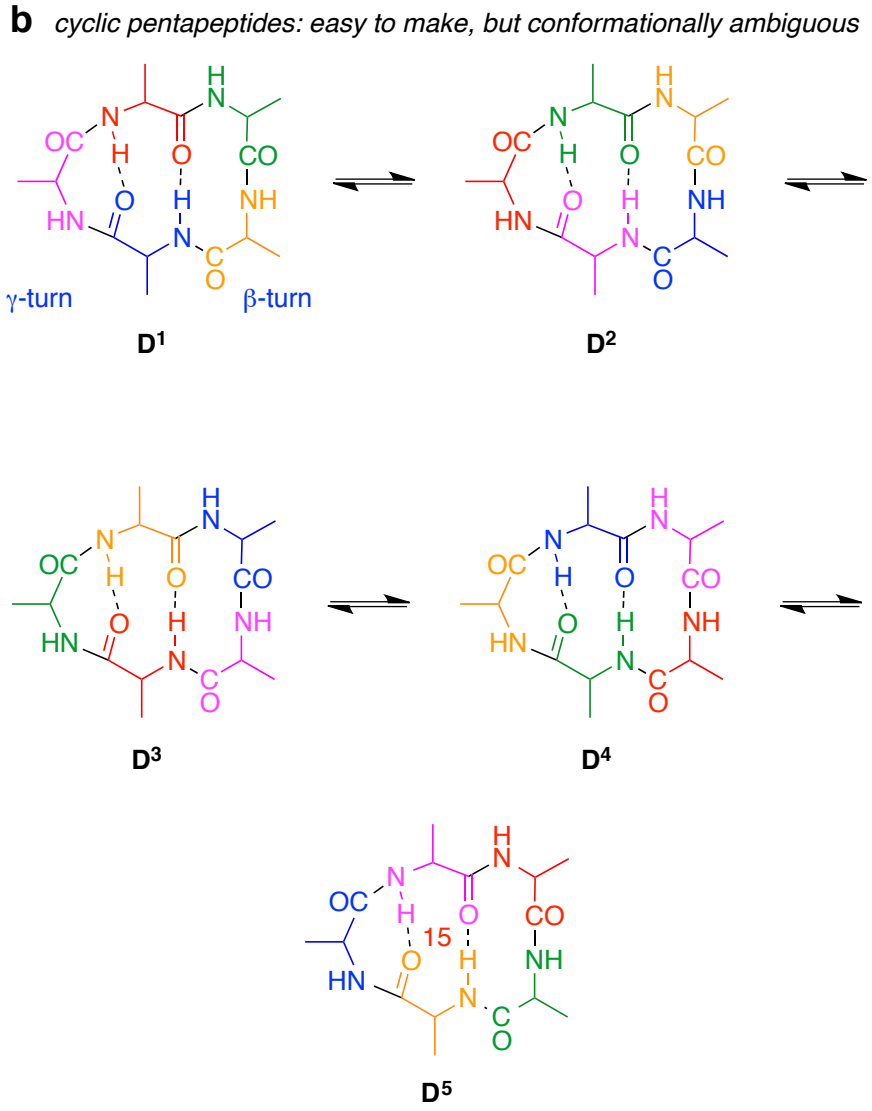
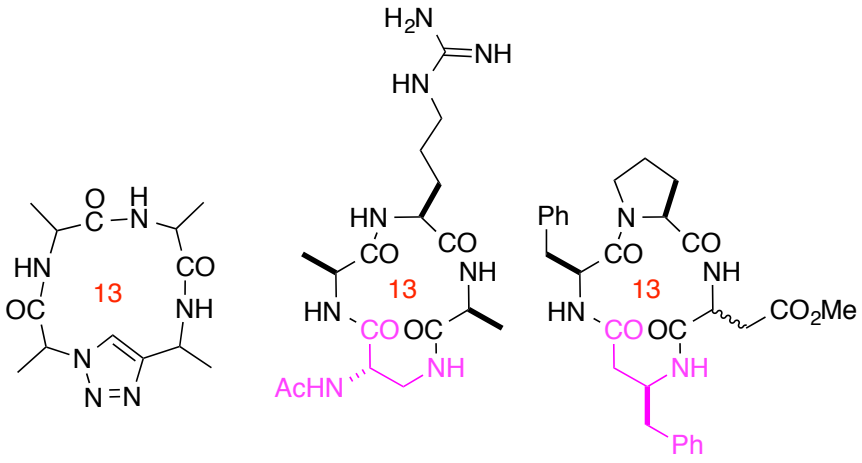


Figure 5.1. Continued.

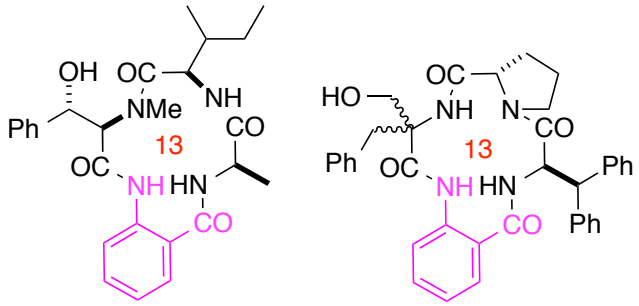
C *some known 13-membered ring cyclic peptidomimetics*



E *Ghadiri et al*

F

G *Fairlie*



H *asperterrestide*

I *synthetic with some antibiotic properties*

Figure 5.1. Continued.

Based on the observations above, there should be favored ring sizes in peptidomimetic design where non-genetically encoded residues replace one amino acid to give conformationally rigid 13- or 14-membered rings. Ghadiri and co-workers, for instance,[161,162] have used copper-mediated azide-alkyne cycloadditions to give the 13-membered rings **E** which were conformationally rigid.[163] In other illustrative work,

Fairlie *et al* substituted β-amino acids into cyclic tetrapeptides and found some 13-membered ring systems, including **F**[164] and **G**[165], that could be prepared efficiently, and were conformationally rigid. However, that same work showed similar, but conformationally *heterogeneous*, 13-membered ring systems.[165]

Anthranilic acid is readily available and more rigid than most other β-amino acids. Several peptidic macrocycles containing anthranilic acid occur in Nature, most where this unit is one of five in a pentapeptide ring.[166-179] There are also numerous examples of similar hexapeptides and higher homologs incorporating anthranilic acid,[166,180-186] a few 10-membered tripeptide derivatives[187-189] and several cyclic systems containing the "*Anth*" residue and another non-encoded amino acid.[190-192] However, 13-membered ring systems containing this ubiquitous residue have been under-explored. Only one natural tetrapeptide **H** that features *Anth* in a 13-membered ring has been discovered,[193] and the only 13-membered cyclic peptide containing *Anth* that has been synthesized is compound **I**, prepared as part of a medicinal chemistry project.[194] To the best of our knowledge, neither **H** nor **I** have been studied in solution to determine their conformational biases.

We hypothesized compounds **1** (Figure 5.2) could be prepared from readily available starting materials, and would be conformationally rigid. This chapter describes how those compounds were made, and the conformational biases of one complete set of enantiomers in this series. In the event the conformations of these molecules were shown to correlate with their chiral amino acid stereochemistries in a logical, easily understood, way that is useful for predicting the preferred shapes of these rigid scaffolds.

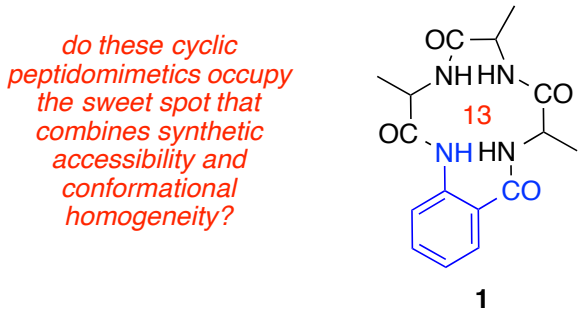


Figure 5.2. The structure of designed constrained cyclic peptide.

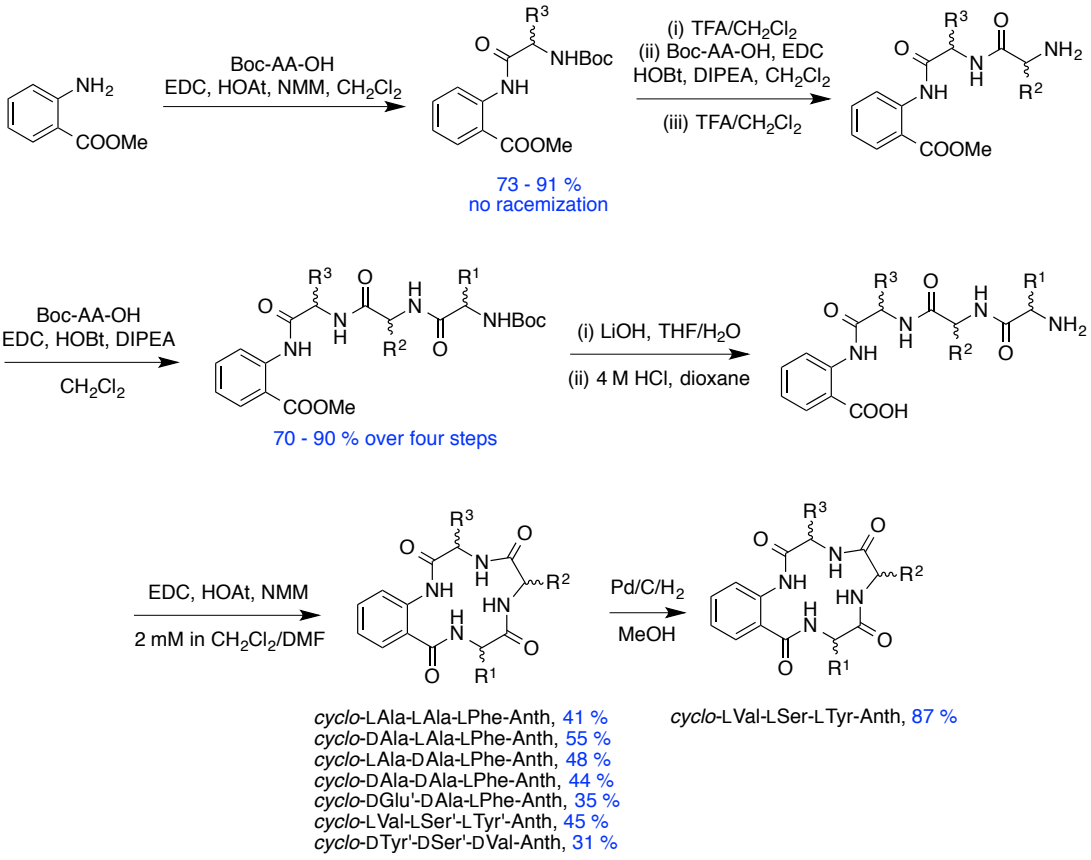
5.2 Syntheses of Cyclic Peptides via Iterative Precipitations

Couplings to anthranilic acid are not facile because the aromatic amine is deactivated via resonance. However, Scheme 5.1 describes how solution-phase syntheses of several compounds were achieved using a large excess of *Anth* and a high concentration of all agents; if high concentrations were not used then epimerization was competitive with product formation. Use of a relatively weak base (*N*-methyl morpholine) and of the superior, though more expensive, carrier agent HOAt,[80] was also beneficial in this step.

Early in this study we realized the physiochemical properties of peptides containing anthranilic acid facilitated their isolation. Thus, coupling three amino acids to the *Anth* unit gave products that precipitated from dichloromethane/hexanes with sufficient purities to use in the next steps. In fact, the only chromatography needed in the "Boc-approach" to the cyclic systems shown in Scheme 5.1 was to isolate the cyclized product. Serine and tyrosine residues in the compounds prepared were protected with benzyl groups. Glutamic acid side-chain protection was achieved using a

tert-butyl ester, which withstands selective deprotection of the *N*-Boc functionality with 4M HCl.[195]

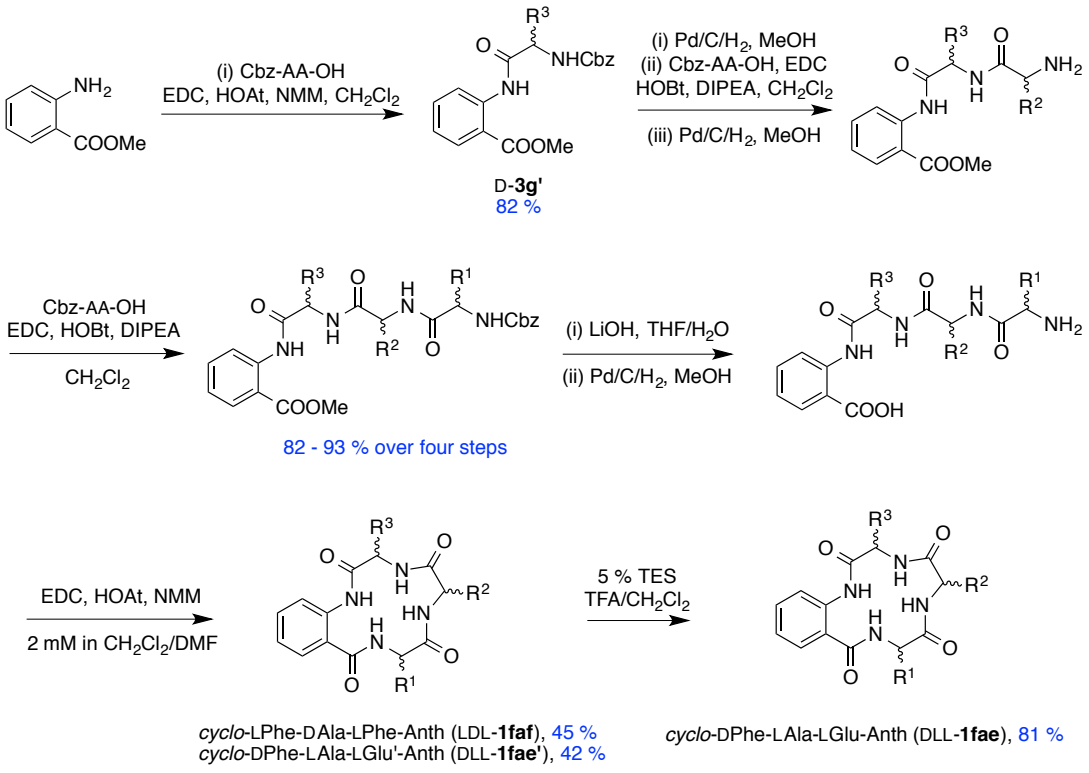
Scheme 5.1. *Boc approach* to products **1**.



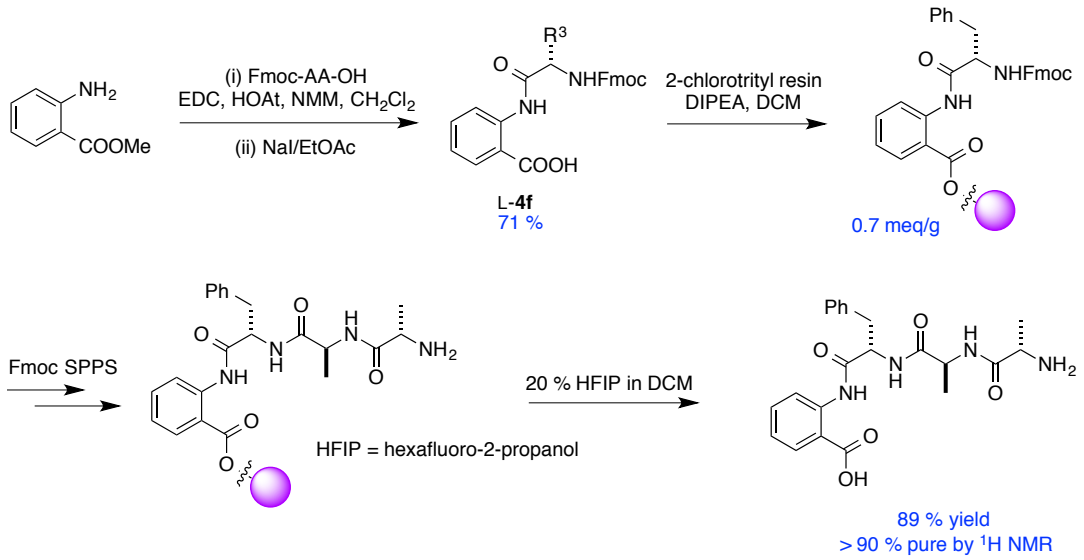
Scheme 5.2 shows a similar solution-phase approach to the same types of products but using Cbz protected amino acids and *Nα*-deprotection via hydrogenolysis. In one of these syntheses the first amino acid added was H-Glu(O^{t}Bu) where successive deprotection under acidic conditions, as in Scheme 1, might have eroded the yield of this

product, but *N*-deprotection via hydrogenolysis circumvented this potential problem. A complementary *solid phase* Fmoc-approach was also established (Scheme 5.3), based on 2-chloro-trityl polystyrene resin and involves cleavage from the support then cyclization in the final step.

Scheme 5.2. *Cbz approach* to products **1**.



Scheme 5.3. *Fmoc approach* to products **1**.



5.3 Conformational Analyses

One goal in this study was to elucidate the intrinsic conformational biases of the stereoisomeric scaffolds **1** with minimal perturbation from side-chain interactions. The ideal system to study might have been the simple peptide **1aaa** (or *cyclo*-AlaAlaAla*Anth*). However, there was insufficient dispersion of the ^{1}H NMR peaks in that particular compound to facilitate convenient conformational analyses. Consequently, we decided to study one enantiomeric series of the compounds **1aaf** (or *cyclo*-AlaAlaPhe*Anth*).

One-dimensional ^{1}H NMR of the **1aaf** series all showed sharp, resolved peaks for the side-chains and for the amide protons. All the amino acid amide *N<u>H</u>* resonances were split into doublets via coupling to the *C<u>H</u>* protons (see below, Table 5.1). These observations imply that the compounds do not equilibrate between different

conformations on the NMR time-scale, and that in solution each exists predominantly in one form. Temperature coefficient data[196,197] and rates of H/D exchange[198,199] were also measured. Neither set of data were particularly informative, except that they indicate there are no "*endo*-cyclic" *H*-bonds, consistent with the conformations deduced from NMR and calculations that are described immediately below. NOE data were collected for all the compounds, and no *cis*-amide bonds were present (Cα*H* – Cα*H* cross-peaks absent).

Table 5.1. Comparison of experimentally observed $^3J_{NH\text{-}\alpha}$ coupling constants with those calculated using the NMR constrained structures simulated in Figure 5.3b.

	Ala(R^1)		Ala(R^2)		Phe (R^3)	
	exp'l[a]	calc.[b]	exp'l[a]	calc.b	exp'l[a]	calc.[b]
LLL	5.0	6.2-6.7	-	7.8-8.3	7.0	6.5-7.8
DLL	7.7	7.6-8.2	8.9	6.6-7.9	7.9	8.0-9.0
LDL	7.2	6.8-8.0	8.8	7.9-8.9	7.7	6.9-8.1
DDL	5.2	5.6-6.3	-	6.6-7.1	5.6	6.6-7.1

[a] Directly from NMR spectra. [b] Based on the structures simulated from the NOE data, and calculated using the Poulson form of the Karplus equation.[200]

Two methods were used to deduce the predominant conformations of compounds **1** in polar solvents. First, the molecules **1aaa** were simulated in a medium of dielectric 46.7 representing DMSO (and 80 representing water, see supporting) using the quenched molecular dynamics (QMD) technique.[75,76] QMD simulations are valuable because they are not bias by NOE data which over-represents some conformations because the NOE effect depends on $1/r^6$ distance relationships.[201] Moreover, QMD thoroughly explores possible local minima in a Boltzmann equilibrium. Second, and independently, NOE

restraints *were* applied in a Macromodel[202] simulation of molecules **1aaf** in dielectric 46.7.

One conformational cluster (maximum RMSD of the Cα – Cβ coordinates 0.5 Å) arose from the QMD simulations of each of the **1aaa** stereoisomers, Figure 5.3a (QMD simulated scaffold conformations are shown in black throughout this chapter). The fact that >1000 conformers (all below 3 kcal•mol^{-1} of the lowest energy one identified) *all* converged to one cluster emphatically indicates conformational rigidity. Comparison of the QMD-generated structures with the NOE data showed they are consistent. Similarly, the MacroModel simulations *with NMR constraints* also gave one predominant conformation for each **1aaf** stereoisomer, Figure 5.3b (MacroModel simulated scaffold conformations involving NOE constraints are grey throughout). None of the $^{3}J_{\text{NH-}\alpha}$ coupling constants were above 9 Hz (see Table 5.1); only couplings >9 Hz are indicative of unambiguous calculated dihedral angles, hence none were used as constraints in the MacroModel calculations.

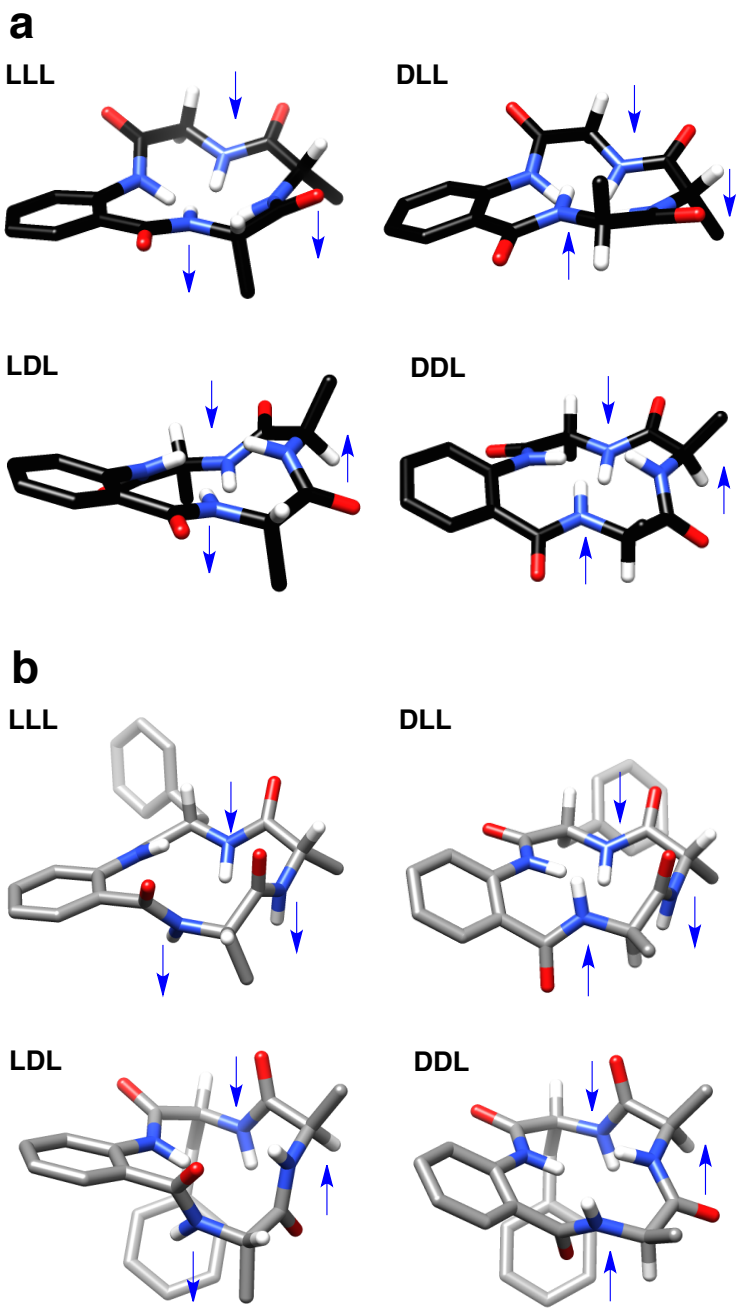


Figure 5.3. Conformations simulated by QMD and NMR conformations. **a** QMD calculations for **1aaa** stereomers *without* NOE constraints; and, **b** MacroModel for **1aaf** stereomers with NOE constraints. The blue arrows approximate the orientation of the NH bond vectors to either pointing up or below the ring system.

Overlays of the conformations generated using the approaches described above, *ie* without and with NOE constraints, showed close agreement (RMSD 0.18 – 0.34 Å for the Cα – Cβ coordinates; Figure 5.4). Moreover, even though the fit of these overlays was based on Cα – Cβ vectors, it is clear that the ring structures also are very similar.

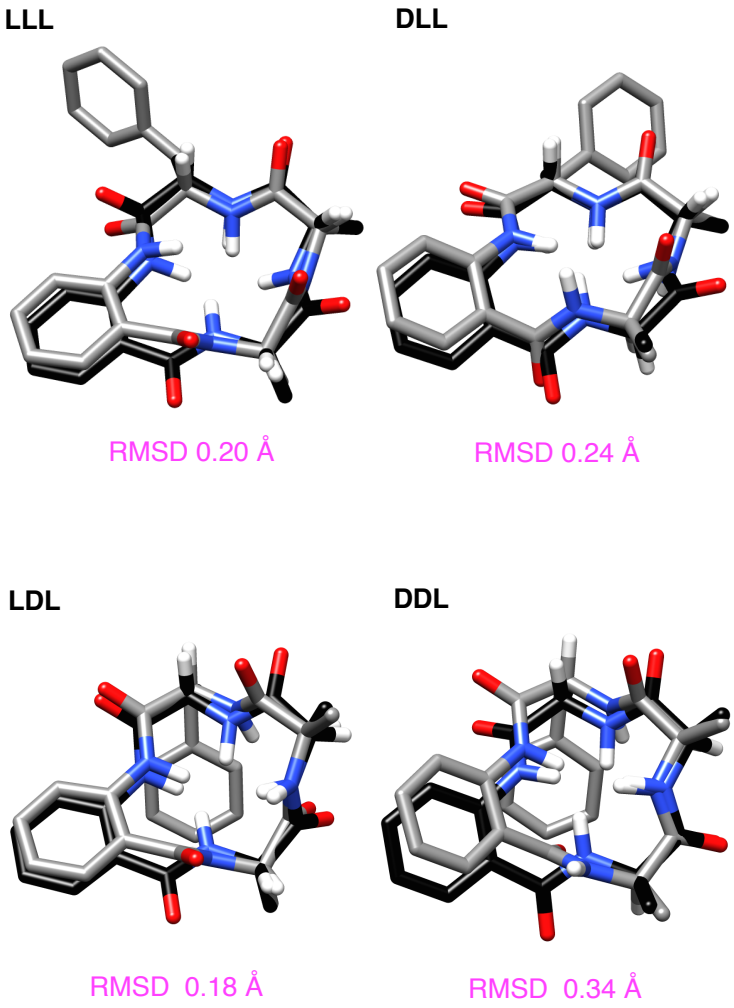


Figure 5.4. Overlays of **1** simulated structures without (black, **1aaa**) and with (grey, **1aaf**) NOE constraints are within 0.18 – 0.34 Å based on 6 coordinates.

Comparisons of favored conformations for stereomers of **1** revealed a simple correlation. When drawn with the ring on the equator, and the *Anth* residue on the West, then L-amino acids point their *N-H* vectors South, while *N-H* vectors for D-amino acids are North-oriented, *irrespective of the other amino acid stereochemistries*. Thus the cyclic 13-membered ring scaffold is constrained to one highly predictable conformation per stereoisomer. This observation held for all the compounds prepared in this study based on similarities in NOE cross-peaks, *ie* for *cyclo*-DGlu'-DAla-LPhe-Anth, *cyclo*-LVal-LSer'-LTyr'-Anth, *cyclo*-DTyr'-DSer'-DVal-Anth, *cyclo*-LVal-LSer-LTyr-Anth, *cyclo*-LPhe-DAla-LPhe-Anth, *cyclo*-DPhe-LAla-LGlu'-Anth, *cyclo*-DPhe-LAla-LGlu-Anth (see Appendix E); those observations imply the conformations are governed by the scaffold and the side-chain variables are less significant.

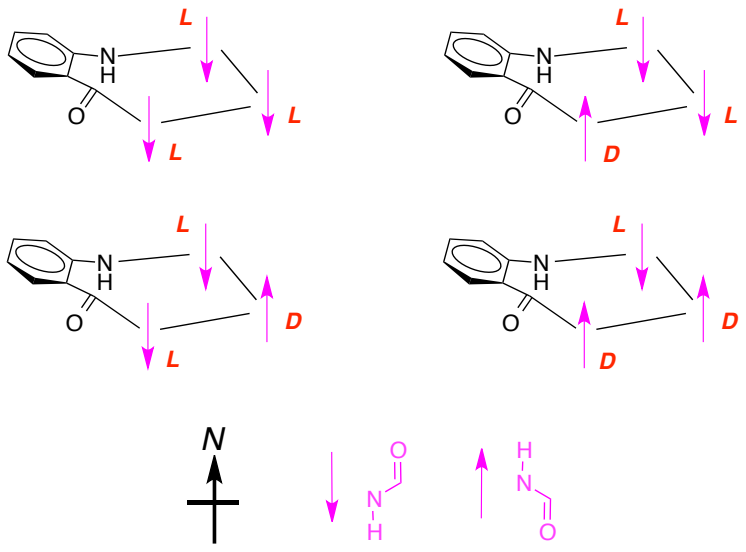


Figure 5.5. "North south orientations" of the amide *N-H* vectors in stereomers of **1** correlate with the amino acid configurations. L-Amino acid *N-H* vectors point South, and while the D-isomers give North-aligned local *N-H* vector orientations, when drawn in the orientations shown.

Table 5.1 compares N*H*-Cα*H* coupling constants directly from NMR spectra with those calculated from the MacroModel simulations involving NOE constraints. In some cases the coupling was obscured, and in several cases the true *J*-values were marginally (by 1.2 Hz at most) outside the calculated range, but the rest were consistent with the values inferred from simulations.

Circular dichroism (CD) data for the **1aaf** stereomers are shown in Figure 5.6 (solid lines). Unsurprisingly, significantly greater molar elipticities were observed for these compounds compared with related linear peptides (dotted lines), indicative of more conformational ordering in the cyclic systems. Moreover, there are logical trends in the data. For instance, the LLL-**1aaf** stereomer (red line) has minima at *ca* 195 and 215, and a maximum at *ca* 230 nm. Substitution of two L-amino acids in LLL-**1aaf** giving DDL-**1aaf** is accompanied by near complete inversion of the CD maxima and minima. The two possible "intermediate" diastereomers, *ie* DLL-**1aaf** and LDL-**1aaf**, in which only one L-amino acid of LLL-**1aaf** is replaced, show shallower peak intensities. The CD spectrum of DLL-**1aaf** (green) is more closely related to that for LLL-**1aaf** (red) than it is to DDL-**1aaf** (purple), whereas for LDL-**1aaf** (blue) the inverse is true; this implies the amino acid opposite the *Anth* residue in the cyclic scaffold has a more profound effect on the molar elipticity.

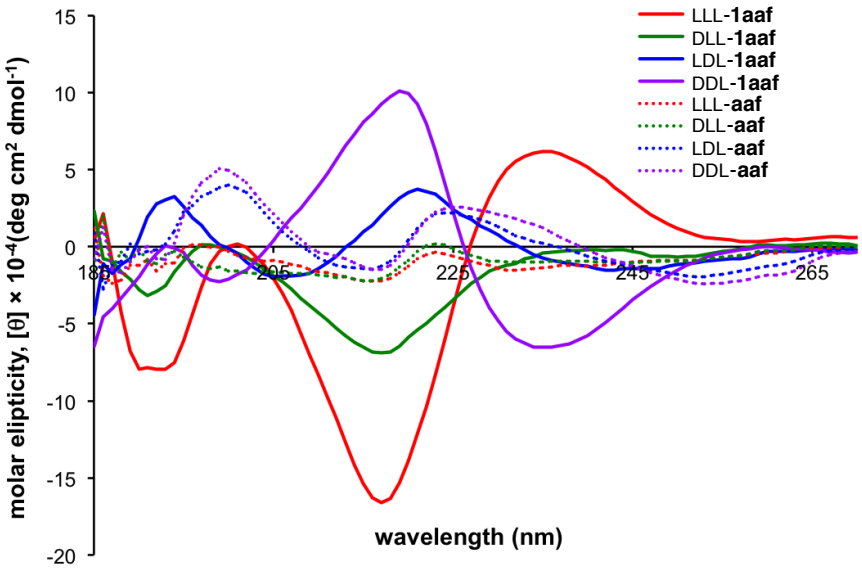


Figure 5.6. CD spectra of compounds **1aaf** (solid lines) and closely related linear peptides (dashed lines).

5.4 Comparisons of the Anth-Cyclic Peptidomimetics with Peptide and Protein Structures

Exploring key orientations on secondary structures (EKOS)[57] facilitates comparison of all the favored QMD simulated conformers with all the common ideal secondary structures based on Cα – Cβ coordinates. For **1aaf** there was only one preferred conformer cluster for each stereomer. Not surprisingly, then, most stereomers matched on only one secondary structure, or none at all, *ie* they are not universal peptidomimetics.[46,47] One apparent exception was LLL-**1aaf** which gave an acceptable fit on γ- and type II β-turn conformations (Figure 5.7); however, γ- and type II β-turn conformations have similar side-chain orientations that overlay well on each other. An

inverse γ-turn matched reasonably well with LDL-**1aaf**, and DLL-**1aaf** overlaid closely with a type I β-turn.

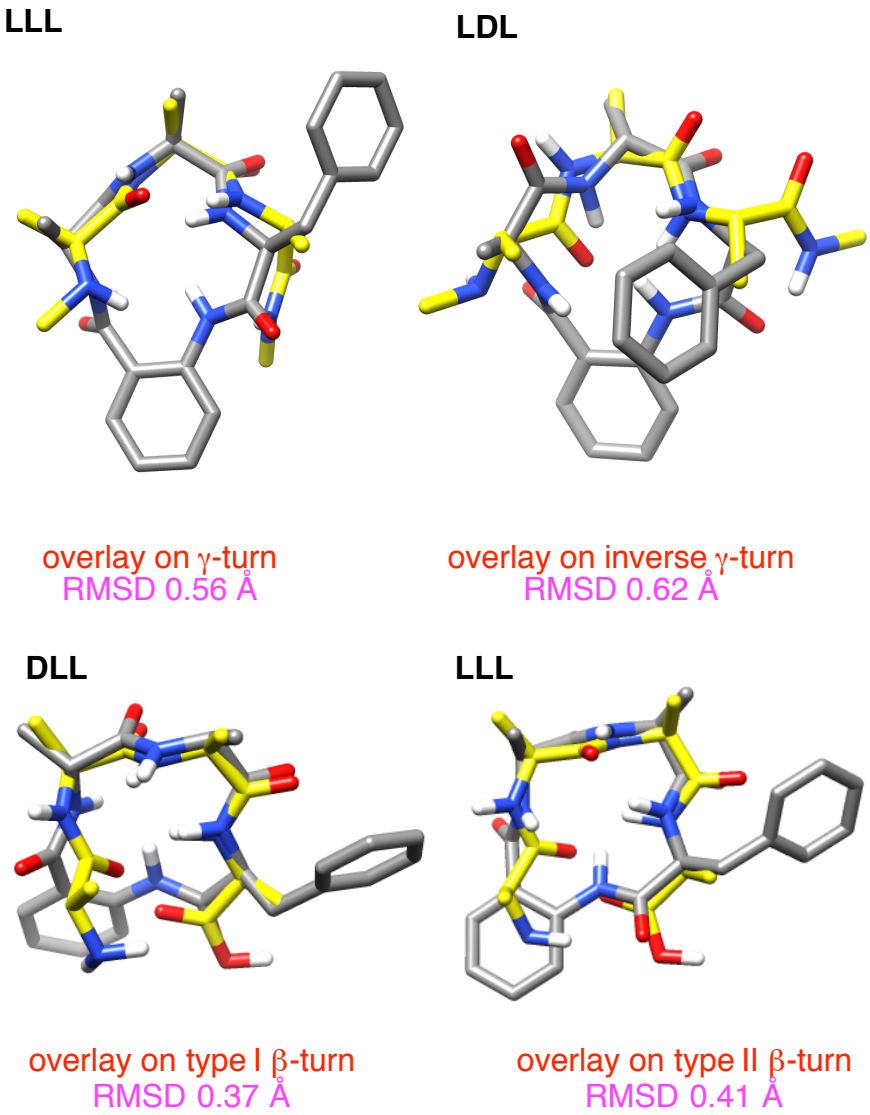


Figure 5.7. Preferred conformations of select stereoisomers of **1aaf** overlaid on γ-, inverse γ-, type I β-, and type II β-turns. RMSD values indicated are for overlay of the side-chain Cα – Cβ vectors.

Whereas, Figure 5.7 depicts preferred overlays of select **1aaf** stereoisomers on ideal secondary structures, the EKO routine[58] facilitates matching QMD generated

structures of **1aaa** with crystallized protein-protein interaction interfaces based on Cα – Cβ coordinates. Thus, preferred conformations of each stereomer were compared with around 160,000 protein-protein interfaces, and each match of RMSD <0.3 Å was analyzed in terms of what secondary structure type at the interface was implicated in the overlay. There were between 106 - 258 matches of <0.3 Å RMSD for each stereomer. The term "no secondary structure" is used here to describe situations in which the region overlaid did not contain any discernable secondary structure. "Turns" refers to a turn of any type (α, β, γ, δ) that has appropriate intrachain hydrogen bonds, while loops and turns without any intra-ring *H*-bonding interactions are classified as "bends". Figure 5.8 shows the distribution of overlays within those categories and some other secondary structures. Thus the preferred **1aaa** conformations are strongly bias towards turns and bends, consistent with the EKOS study presented in Figure 5.7.

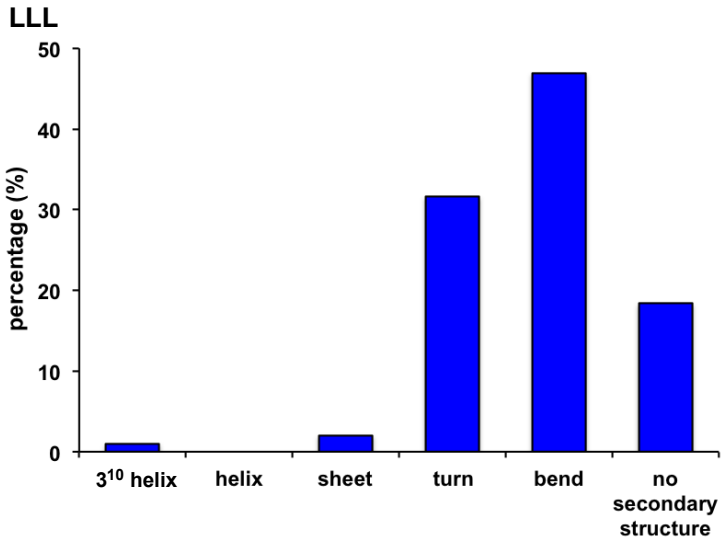


Figure 5.8. Distribution of best overlays on PPI interface segments with respect to secondary structure for the stereomers featured in Figure 5.7.

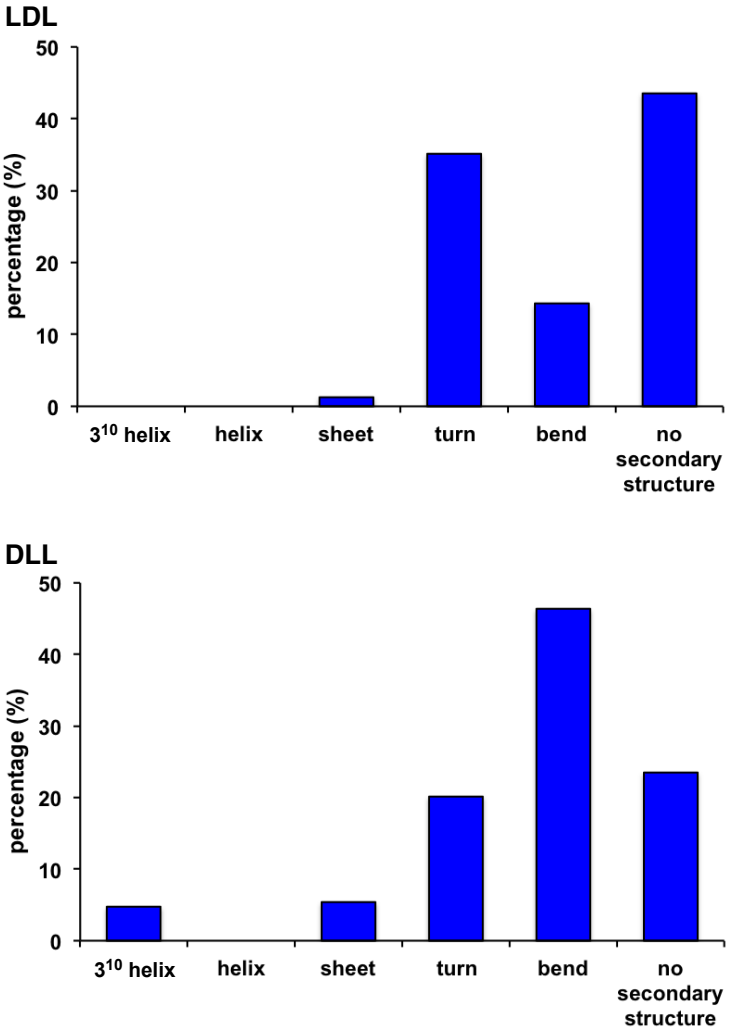


Figure 5.8. Continued.

Finer detail of the secondary structure types that were overlaid in the EKO analysis of PPIs in the PDB is indicated by ϕ,ψ-angles, indicative of the type of secondary structure implicated. Figure 5.9 shows select data for the stereomers featured in Figure 5.7. Thus, the number of occurrences where LLL-**1aaa**, LDL-**1aaa**, and DLL-

1aaa overlaid closely with γ-turns, type I β-turns, inverse γ- and type II β-turn conformations are consistent with the favored conformations predicted in Figure 5.7.

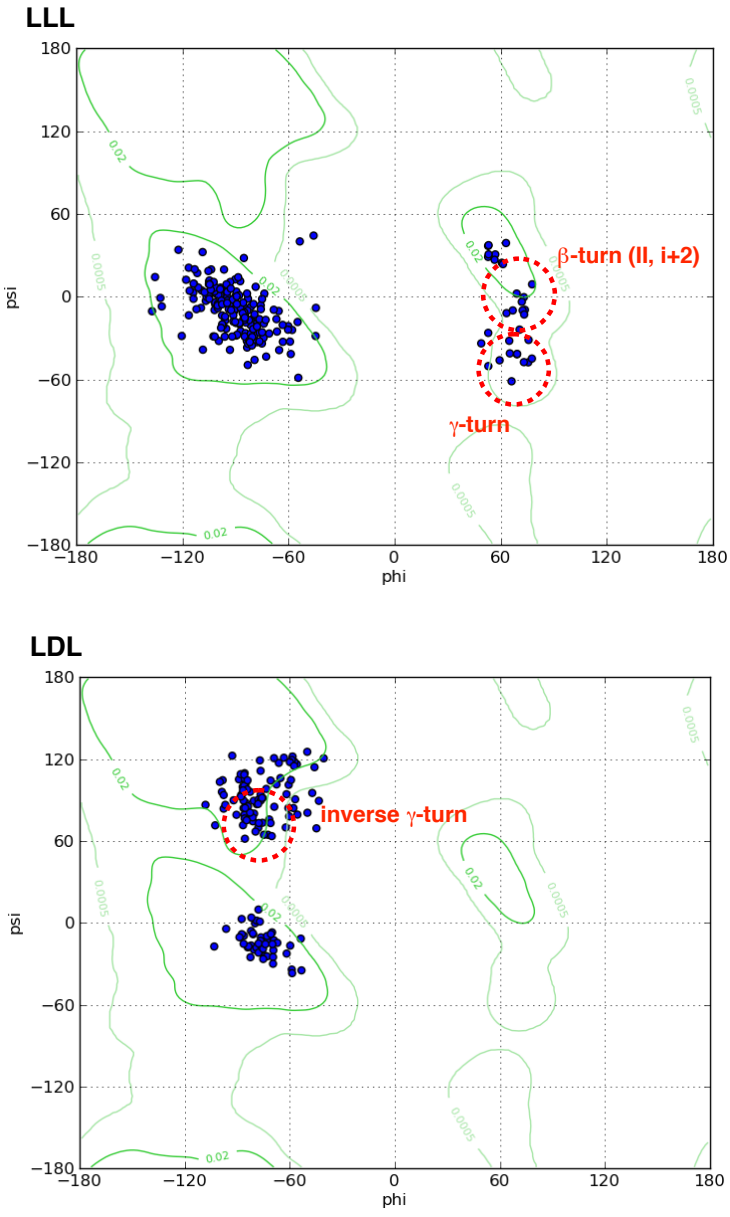


Figure 5.9. Each dot on these plots is associated with a ϕ,ψ-bond vector of a protein interface region that overlaid closely with a preferred conformer of the **1aaa** stereomer indicated.

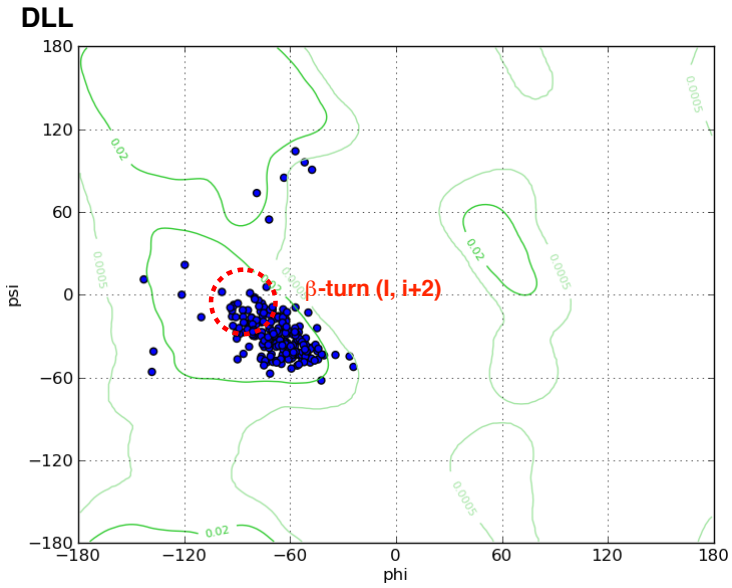


Figure 5.9. Continued.

5.5 Physiochemical Properties

HPLC was used to monitor the stability of LLL-**1aaf** under aqueous conditions of pH 2,7, and 12, and, in separate experiments, the mixture of proteases called *Pronase*[203] that is used to extensively hydrolyze proteins in proteomics studies (any amide between two hydrophobic residues could be cleaved by this enzyme mixture).

Table 5.2. Stabilities of LLL-**1aaf** under different pH conditions**.**

	conditions	$t_{1/2}$ / h
pH stability of LLL-**1aaf**	pH 2.0[a]	>500
	pH 12.0[b]	240
	pH 7.4[c]	>500
LLL-**1aaf**	Pronase[c]	no cleavage after 12 h
linear LLL-**aaf**	Pronase[c]	1.5

[a] 10 mM HCl, [b] 10 mM NaOH, [c] PBS buffer. 20 % MeOH was added in all cases to increase solubilities.

Figure 5.10 shows data calculated (QikProp)[204,205] for the polar surface area (PSA) and cell permeability of LLL-**1aaf**. The implications of this data are that LLL-**1aaf** has less PSA and a higher tendency to permeate into cells than closely related linear peptide controls. Polar surface areas <140 $Å^2$ are generally preferable for cell permeability.[206] Similarly, compounds with predicted PCaco-2 permeability rates >20 nm/s are usually cell permeable; data calculated for LLL-**1aaf** exceeds both expectations, and predict cell permeability.

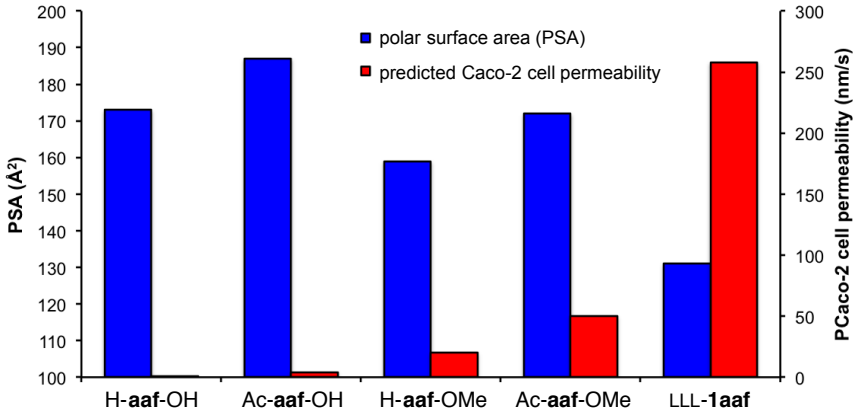


Figure 5.10. Comparison of PSA (blue) and predicted Caco-2 cell permeabilities (red) for linear peptides based on **aaf** and a featured cyclic molecule **1aaf**.

5.6 Conclusions

The *Anth* containing peptides featured in this study have a greater tendency towards conformational homogeneity than systems based on other β-amino acids related to **F** and **G**.[165] In view of this, and other observations outlined in the introduction, we suggest 13-membered systems *based on the Anth residue* are at a favorable point on the crossroads between ease of synthesis and conformational rigidity.

Predictability of conformational preferences for a set of cyclic peptidomimetics is an attractive feature of the systems studied here. It is possible that different side-chains to the ones in **1aaf** could perturb the predicted conformations, and this study does not encompass the special case of systems that contain Pro- or Gly-. Nevertheless, the data above clearly indicates strong intrinsic conformational biases of the scaffolds based on other amino acids. Moreover, stereochemistries of the scaffolds can be manipulated so that they orient their side-chains in ways that resemble the same orientation for regular and inverse γ-turns, and for the two most common β-turns (types I and II).

Consistent with this, conformations of **1aaa** stereomers have a pronounced tendency to orient side-chains in ways that some turns and bends do at PPI interfaces. Overall, while tens of naturally occurring cyclic peptides containing *Anth* have been discovered, the potential of this simple residue in designs of conformationally rigid protein interface mimics has hitherto been under-appreciated.

CHAPTER VI

EVALUATING MINIMALIST MIMICS WITH EXPLORING KEY ORIENTATIONS ON SECONDARY STRUCTURES*

6.1 Introduction

Semi-rigid scaffolds that express amino acid side-chains are interesting chemotypes because these "minimalist mimics"[47] of secondary structure motifs have the potential to disrupt protein-protein interactions (PPIs). They have this potential because PPI interfaces are dominated by side-chain to side-chain contacts[48] hence semi-rigid small molecules that project side-chains in similar orientations to one protein component might competitively interact with the protein-binding partner. Minimalist mimics do not have peptidic polyamide backbones so they are not degraded by proteases and tend to have improved cell- and oral-bioavailabilities relative to peptides.[207]

Chronologically, scaffolds **1**,[208-210] **2**,[211] **3**,[212-214] **4**,[215,216] **5**,[217] **6**,[218] **7**,[45] and **8**[219] (Figure 6.1) were reported as mimics of helical secondary structures, and are typical of ones in the literature.[44,220] Our interest in the concept of "universal mimics", wherein several secondary structures are represented in one conformational ensemble,[46] led us to wonder if scaffolds that had been described only as α-helical mimics might also be able to access conformers that match other secondary structures. None of the α-helical

---

*Reproduced in part from "Evaluating Minimalist Mimics by Exploring Key Orientations on Secondary Structures (EKOS)", Dongyue Xin, Eunhwa Ko, Lisa M. Perez, Thomas R. Ioerger and Kevin Burgess, *Org. Biomol. Chem.* **2013**, *11*, 7789-7801.

mimics **1** – **8** have been claimed to orient side-chains in ways that resemble the less common 3_{10}- and π-helical motifs, but it seemed possible that they could at least do this. Moreover, there is another important issue that it is convenient to refer to here as *side-chain correspondence*; by this we mean the particular side-chains in a secondary structure that best overlay on an accessible conformation of a minimalist mimic. For instance, a scaffold that rigidly presents side-chains in an *i*, *i*+4, *i*+7 orientation has that *side-chain correspondence*, whereas it might be unsuitable as a mimic of *i*, *i*+3, *i*+4 helical motifs because this is a different *side-chain correspondence*. When evaluating molecules to perturb different protein-protein interfaces it is at least highly desirable, and probably essential, to have a selection of mimics that cover a range of possible side-chain correspondences, and it is important to understand what these are for a given mimic. Of course, the community working in this area is well aware of the need to use the correct side-chain correspondences, but there is no rigorous, systematic approach to assessing what they are for a particular minimalist mimic.

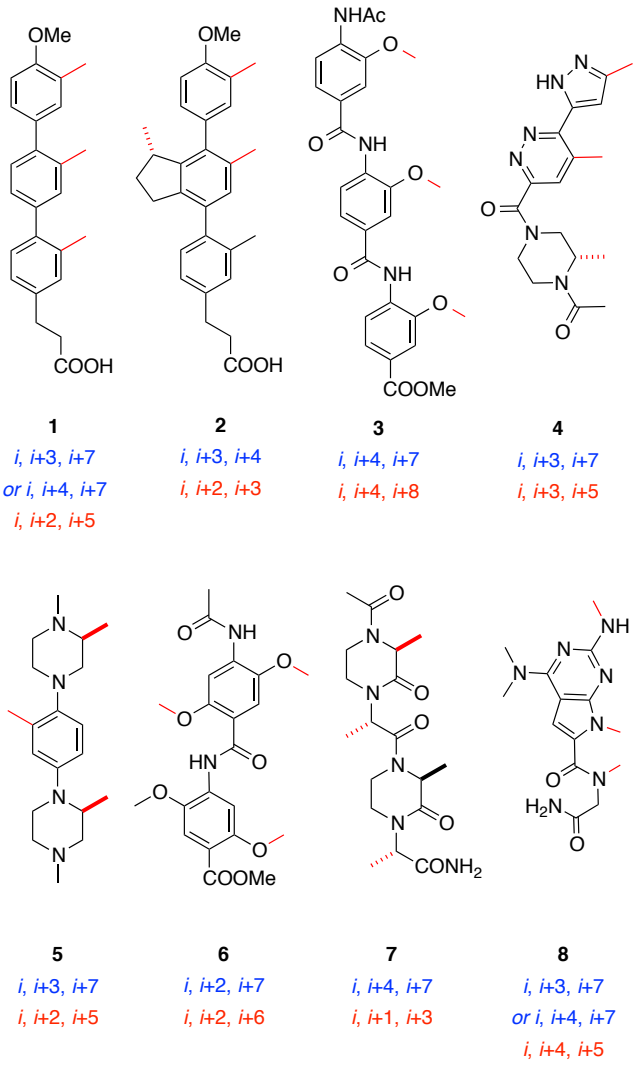


Figure 6.1. Helical mimics featured in this chapter.

After a side-chain correspondence has been determined, there needs to be a standard parameter to gauge how each minimalist helical mimics fits on that particular side-chain combination. For instance, if two mimics orient *i*, *i*+3, *i*+7 side-chains in an ideal α-helical conformation, which one does it most closely? What are the best

minimalist mimics to present side-chains in *i*, *i*+3, *i*+4 and *i*, *i*+4, *i*+7 orientations corresponding to an ideal 3_{10} helix? Questions like these are difficult to answer using the approaches that have been employed in the literature so far; currently there is no widely accepted method to evaluate the bias of a given scaffold toward a particular conformation. One objective of this chapter is to propose a strategy that is applicable to this issue.

Several *experimental* methods have been applied in the context of elaborating equilibrating conformations of minimalist mimics in solution, but none of them, or any combination of them, are suitable for determining the side-chain correspondences and goodness of fit of helical mimics. Broadly speaking these are methods to observe the mimics in solution (*eg* NMR and CD), and crystallography.

Arora's study of helical mimic **7**[45] illustrates what direct spectroscopic evidence can be collected to characterize equilibrating conformational states of a minimalist mimic in solution. Circular dichroism (CD) of the molecules **7** in methanol and in acetonitrile had a similar shape to ones for peptide helices, but the minima were red-shifted by 10 nm. NOE experiments were conducted on one compound (actually, in $CDCl_3$), and cross peaks that correspond to the trans-isomer, but not the cis-one, were observed. However, in solution, minimalist mimics exist as rapidly equilibrating conformational states for which CD and NOE data are averaged. In NMR, NOE cross peaks over-represent close contacts relative to distal ones because of the inverse six-power distance relationship; consequently, poorly populated states can appear to be more abundant than prevalent ones. Moreover, some NOE cross peaks for molecules of

approximately this molecular mass can be vanishingly small, so absence of cross peaks does not prove that the corresponding conformation is unpopulated; ROESY spectra can help with this, but they have to be carefully calibrated to the properties of the molecule.[221,222] In any case, minimalist mimics definitely have multiple conformations, so it is very difficult to unravel which sets of NOE crosspeaks, or CD elipticities, are associated with individual ones. Consequently, none of the techniques often used[53,214,223] give detailed information on the ensemble of preferred conformations in solution.

Researchers often turn to crystallography to elucidate conformations of minimalist mimics. Crystallography reveals only a few conformations at best, and these may not be representative of favored ones in solution because of crystal packing forces. Nevertheless, crystallography may provide circumstantial evidence that desired solution states can be accessed, and this type of conclusion is most convincing when molecules crystallize in a conformation that is anticipated to be strongly preferred in solution. This is true of the benzamide mimics **3** and **6** where H-bonding between the amide-N<u>H</u> and the 2-methoxy substituent is expected to disfavor rotation about the aryl-NH bond, and X-ray structures show this.[224,225] Similarly, an X-ray crystal structure of one of Hamilton's terphenyls was obtained to reveal a relevant solid state conformation.[210] However, minimalist mimics sometimes crystallize in conformations that are not relevant for secondary structure mimicry. Single crystal X-ray studies reported in a recent contribution from Hamilton, for example, describe how a scaffold that is putatively a sheet mimic in solution crystallizes in some other conformation.[55] In

summary, crystallography does not characterize rapidly equilibrating conformational ensembles of minimalist mimics in solution.

Given that there seems to be no experimental strategy to characterize equilibrating conformations of minimalist mimics in solution, researchers often turn to calculations. Hamiliton's terphenyls **1** were originally conceived to match *i*, *i*+3, *i*+7 side-chains on an ideal α-helix (though he has used it as an *i*, *i*+4, *i*+7 mimic on some actual helices in PPIs).[208,210] To validate that design his group used Still's Macromodel[202] to simulate some of the preferred trimethyl-substituted terphenyl conformers, and focused on an accessible state that had similar angular projections to the *i*, *i*+3, *i*+7 side-chains of an ideal (all-Ala) α-helix. Incidentally, it is relevant to what follows in this work that they observed their simulated structure had 4 – 25 % shorter distances than was optimal, and that the root mean square deviation (RMSD) of the overlay based on Cα - Cβ coordinates was 0.85 Å.[210] Similarly, Arora's hypothesis for α-helical mimicry by **7** is based on modeling of that scaffold using Macromodel (MMFF force field in chloroform). In both the Arora and Hamilton work, the existence of one preferred helical conformer supports the assertion that the scaffold is an α-helical mimic,[226] but this does not comment on the other equilibrating conformations in solution. In other work, researchers have used molecular dynamics simulations on mimics **3** to determine if and how these molecules might bind a groove in the p53 protein that recognizes a helical motif on MDM2,[227] and for the free mimic in water.[228] These studies were not performed with the objective of simulating the ensemble of preferred conformers but instead were used to search for one desired conformer. All

these approaches are intended to find particular conformations of putative helical mimics. Such computational strategies sample multiple conformations, but only select certain favored states for analyses.

Near the beginning of this introduction, three important parameters were outlined for evaluating how minimalist mimics can match ideal secondary structures. Essentially these involve considering every accessible member of a representative conformational ensemble for: (i) overlay of conformers on various secondary structures; (ii) side-chain correspondences; and, (iii) goodness of fit. Conveniently, scaffolds **1** – **8** are relatively rigid, so the calculated bond lengths and angles for each particular conformation are likely to be reliable relative to more complex and flexible structures (*eg* peptides); in other words, simulated conformers of these molecules tend to be quite realistic. However, it is important to recognize the importance of generating and considering a large set of conformers for these molecules; this is because conformers that are quite similar can project side-chains in significantly different ways. For example, sampling a 360° *aryl-aryl* rotation in terphenyls **1** in 1° increments gives a continuum of 360 states, most of which are accessible at room temperature in solution. Within that group of conformers, a difference of only a few degrees in the torsion angle will give significantly different side-chain orientations. Consequently, for the purposes of this work, where most of the accessible side-chain orientations should be considered, it is important to avoid computational approaches that minimize and cluster accessible conformations into families representing local minima. Methods based on Monte Carlo or molecular dynamics with simulated annealing would be inappropriate here. In summary,

minimizing routines that cause conformations to converge on local minima are unsuitable for simulating conformational *ensembles* of minimalist mimics.

A representative conformational ensemble for a minimalist mimic might involve hundreds of conformations, all of which are significantly populated, *ie* within, for instance, 3 kcal/mol of the minimum energy conformer located. It is necessary to match each of those conformations on every three amino acid side-chain combination in every common secondary structure to characterize how a minimalist mimics fits, *ie* to evaluate the three parameters that reveal how minimalist mimics can match ideal secondary structures; this is a data mining problem rather than a modeling issue. It requires specialized algorithms similar to those we already developed for Exploring Key Orientations (EKO) in protein-protein interactions.[58]

On the basis of the considerations outlined above, we devised a strategy to evaluate how accessible conformations of minimalist mimics match a collection of idealized secondary structures. This strategy consist of the following steps:

i. use quenched molecular dynamics (QMD) to generate an ensemble consisting of hundreds of accessible, low energy conformers of the mimics;
ii. represent each of these as a set of Cα and Cβ coordinates corresponding to three amino acid side-chains displayed by the scaffolds;
iii. represent each combination of three side-chains in each ideal secondary structure as a set of Cα and Cβ coordinates corresponding to three amino acid side-chains displayed by the scaffolds; and,
iv. overlay Cα and Cβ coordinates of all the conformers on all the sets of side-chain "triads" in the ideal secondary structures and express their goodness of fit in terms of root mean squared deviation (RMSD, Å) for each overlay.

The strategy outlined above is different to the EKO process[58] which explores key orientations at protein-protein interfaces and is not concerned with secondary structure classifications. However, the strategy outlined here is similarly motivated and facilitated by the process of data mining a conformational ensemble on target structures, so we refer to it as *Exploring Key Orientations on Secondary structures* (EKOS). To the best of our knowledge, it is fundamentally different to any computational approach that has been applied to evaluate minimalist mimics to date. EKOS has the considerable advantage that every accessible conformer in large conformational ensemble is evaluated on every side-chain triad in every secondary structure, and the results are systematically ranked in terms of goodness of fit.

In this chapter, we focus on the α-helical mimics **1** – **8** using EKOS to enable quantitative evaluations of solution state structures that are not conveniently possible via spectroscopy or other methods.[229] Smith's β-sheet scaffold **9**,[230] and an interface mimic **10** developed on our laboratory[62] were used as "controls" in this study. (Figure 6.2) Scaffold **10** has already been shown to analog several secondary structures with a bias towards β-sheets and, relatively speaking, no notable inclination to mimic helical structures.

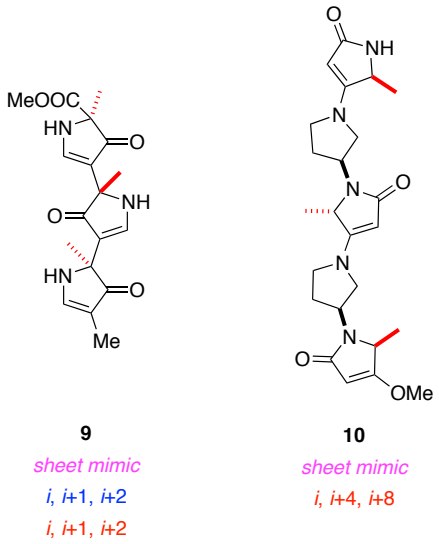


Figure 6.2. Smith's β-sheet **9** and oligo-pyrrolidine-pyrrolidinone **10**.

Throughout these discussions it is important to remember that helical mimicry is only a means to an end: to use these compounds to displace a protein or peptide that has a helical conformation at a PPI interface. In specific applications of this kind, mimicry of ideal helices is less important than matching the actual helical structure at the interface, which can be distorted and non-ideal. Consequently, even though this study is primarily about evaluating *ideal* helical mimics in general, we have also used EKO to compare the accessible conformations of the helical mimics with helical structures in some of the PPIs that have been perturbed using these mimics. The objective of that part of the study was to ascertain how well scaffolds **1** – **8** may fit on some well-studied helices in PPIs and compare this with the data for ideal helical mimics.

6.2 Bases for Comparison: Ideal Secondary Structures and How to Compare Them

Secondary structures at protein-protein interfaces are rarely ideal. α-Helices, for instance, can be stretched, compressed, bent, kinked, and partially unwound. In the minimalist mimic area, the most direct approach to designing a scaffold to specifically perturb a particular PPI is to match the secondary structures at that interface, including their distortions and other peculiarities. However, it is often logical to begin the process of mimicking a real helical motif with the most closely related *ideal* helical mimic. Recognizing this, a significant proportion of the research community who work on minimalist helical mimics base their designs on ideal secondary structures without stating a target PPI; this strategy affords generally interesting data because they mimic *ideal* secondary structures representative of the most common states found in proteins. For the reasons stated above, we decided ideal secondary structures should be the focus of our analysis here. However, the end of this chapter relates the findings to some helical mimics that have been shown to perturb specific PPIs.

Amino acid *C*α and *C*β coordinates are the best simple method for defining side-chain orientations. This is because setting a *C*α - *C*β vector excludes many orientations of the "downstream" side-chain bonds; it would be *inappropriate* to use other side-chain vectors because *C*β - *C*γ linkages and beyond are much less constrained. Overlaying *C*α and *C*β side-chain coordinates may be used to access goodness of fit of a conformational state on a secondary structure in terms of RMSDs. This procedure has been used when applying Bartlett's CAVEAT algorithm,[231,232] Hamilton used it on his mimics of secondary structures,[55,210] and we have used it extensively too.[46,58,62]

6.3 Method for Comparisons of Scaffold Conformations with Ideal Secondary Structures

Preferred conformations of scaffolds that express only *C*α and *C*β atoms, *ie* methyl substituted ones, show how the molecular core is biased to project amino acid side-chains. Any conformational state of all-methyl-substituted compounds may be approximated to three $C\alpha - C\beta$ vectors or six ($3 \times C\alpha + 3 \times C\beta$) coordinates to represent how that state projects side-chains. Thus each conformation can be described in terms of irregular prismatic shapes formed by joining these coordinates (Figure 6.3). Similarly, any set of three side-chains in an ideal secondary structure may be described in the same way.

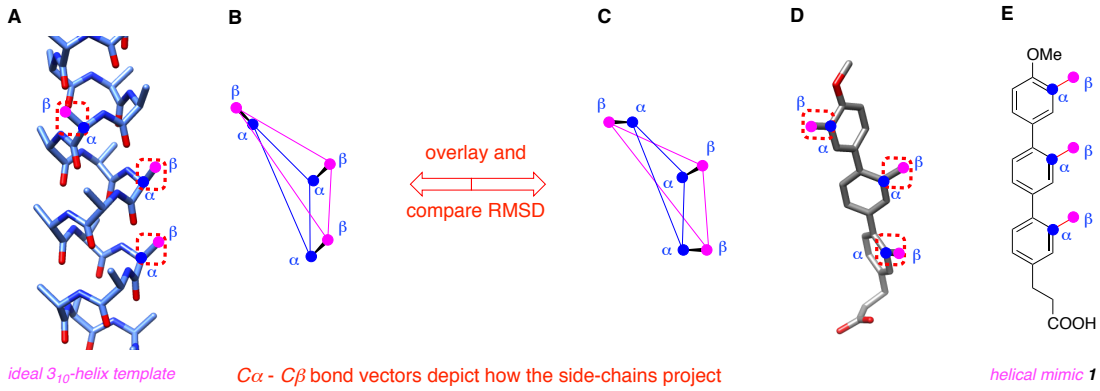


Figure 6.3. Side-chain matching on secondary structures based on 3 x ($C\alpha - C\beta$) coordinates.

Conformational populations are determined by a combination of kinetic and thermodynamic issues. It is only necessary to consider kinetic effects for semi-rigid small molecules if they have restricted rotation at ambient temperature that prevents

equilibration between conformers; this is not the case for scaffolds **1** – **8**. To evaluate the thermodynamically favorable conformational states, we used Pettit's QMD procedure.[75,76] Specifically, the scaffolds were minimized (molecular mechanics) then subjected to molecular dynamics at high temperature (1000 K) for 600 ps; 600 conformational states were thereby recorded during this run (*ie* one every 1 ps). These conformers were then minimized via molecular mechanics and all the simulated conformers within 3 kcal/mol of the most stable identified were considered. Table 6.1 shows that in this strategy the number of accessible conformers generated for each mimic was 282 or more, implying that motions about each significant degree of freedom were explored in small increments by this conformer set. Thus all the accessible conformers generated in this way for each scaffold were systematically overlaid on every combination of side-chains in idealized secondary structures using procedures also used for the Exploring Key Orientations (EKO) approach as described.[58]

Table 6.1. Number of conformers below 3.0 kcal•mol^{-1} for each mimic from QMD.

mimic	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
number of conformers	600	599	567	469	282	507	299	421	600	490

The prime objective of this work was to ascertain the bias of semi-rigid scaffolds **1** – **10** on secondary structures. Consequently, the medium used for the simulations was a constant (featureless) one of dielectric 80, corresponding to an aqueous environment. Explicit water molecules were not used because they would be displaced as the small

molecule begins to interact with the protein-binding partner. Similarly, interactions between specific extended side-chains (*eg* Glu-Lys salt bridges) also tend to be nullified if these are interface side-chains that dock with a protein-binding partner. Assessing all the various side-chain to side-chain interactions that could occur therefore is unimportant for this method, besides being impractical. What is important is the conformational bias of semi-rigid scaffolds **1** – **10** with methyl side-chains, *ie* ones that reveal the intrinsic biases of the scaffolds.

Mimics **2**, **6** and **7** have *four*, not three, side-chains on the scaffold. For **2** the three side-chains highlighted in red (Figure 6.1) were considered because these have different orientations to those in the terphenyl **1**. In **6** and **7** the side-chains chosen were two terminals and one internal because we estimated that combinations of this kind are most likely to resemble those on one face of a helix. Selection of those side-chains for **7** corresponds to the ones Arora originally used to generate overlays.

Six common ideal secondary structures were chosen for the overlay process (3_{10}-, α-, and π-helices; β-strands; parallel- and antiparallel β-sheets). Templates for ideal structures were obtained from Discovery Studio 2.5 (3_{10}-, α- and π-helices, and β-strands) and modified β-sheet builder (www-lbit.iro.umontreal.ca/bBuilder/index.html; parallel- and antiparallel-β-sheet and sheet/turn/sheets). Strand-turn-strand structures were also included, even though these are closely related to antiparallel β-sheets, because mimics **1** – **10** have extended structures that can simultaneously overlap with both the β-turn and β-sheet regions. Figure 6.4 illustrates this, and is also intended to

show that mimics can achieve optimal goodness-of-fit by overlaying on one strand or by "lying across" two or more regions in a sheet.

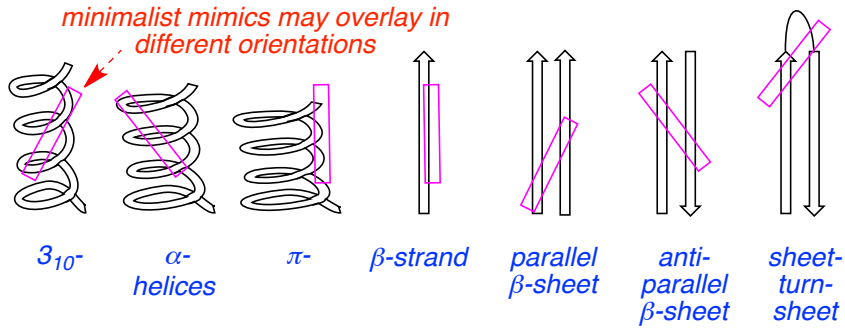


Figure 6.4. Overlays considered in this chapter place mimics on secondary structures in any orientation.

6.4 Comparisons of Scaffold Conformations with Ideal Secondary Structures

Table 6.2 summarizes the lowest RMSDs obtained for fitting all the selected conformers for each mimic on ideal secondary structures. These data are color coded to enable rapid evaluation of trends, but no absolute significance to the color distinctions is implied. The one case that corresponded to an excellent fit is shaded in red, very good ones are shown in yellow, good ones are shown in green, and any "best-fit" with an RMSD of more than 0.70 Å is not shaded. To the best of our knowledge, there are no prior literature reports of minimalist mimics overlaid on α-helical motifs with RMSDs less than 0.70 Å, so these arbitrary delineations of "excellent", "very good", and "good" fits are relatively stringent.

Data in Table 6.2 are remarkable in several ways. As a whole they indicate some putative α-helical mimics actually have conformers that resemble β-strands, parallel and

antiparallel β-sheets, and/or sheet-turn-sheets. Mimic **8** had no conformers that match any of the three helical types at RMSD 0.70 Å or less, but did have ones that fitted well on extended, sheet-related, conformations. Similarly, mimic **7** was shown to be a significantly better strand-turn-strand analog than it was for any of the helices, and the only helical structure that matched well was the rarer π-form. Hamiliton's terphenyl mimic **1** gives better matches on 3_{10}- and π-helices than on the α-form, and overlaid unexpectedly well on a sheet-turn-sheet. In fact, the only very good α-helical mimic in the series was oligobenzamide **3**. Based on this analysis, scaffolds **3** and **4** appear to have the most potential as "universal mimics"[46] since they gave conformers that fit all six secondary structures well. However, this comparison is not even because it treats scaffolds **3**, **6**, and **8** in a *different* way to the others, for the reasons described below.

Table 6.2. Matching preferred conformations of scaffolds on ideal secondary structures[a].

	3_{10}-helix	α-helix	π-helix	β-strand	β-sheet (parallel)	β-sheet (anti-parallel)	sheet-turn-sheet
1	0.65	0.71	0.59	-	-	-	0.58
2	0.39	0.59	0.78	-	0.89	0.81	0.84
3	0.30	0.31	0.47	0.70	0.58	0.43	0.43
4	0.36	0.51	0.55	0.57	0.60	0.59	0.59
5	0.60	0.51	0.55	-	-	-	0.54
6	0.60	0.61	0.57	0.96	0.93	0.90	0.51
7	0.81	0.72	0.51	1.05	0.89	0.95	0.36
8	0.74	0.80	0.85	-	0.61	-	0.61
9	0.82	0.79	0.91	0.65	0.41	0.69	0.44
10	0.58	0.78	0.63	0.73	0.42	0.55	0.46

[a] RMSD ≤ 0.30 Å, red background; 0.31 - 0.50 Å, yellow; 0.51 - 0.70 Å, green. RMSDs of over 1.1 Å are not shown.

The bonds highlighted in red Figure 6.1 are the ones that the original researchers used to overlay with *C*α - *C*β vectors. However, the highlighted bonds for mimics **1**, **2**, **4**, **5**, **7**, **9** and **10** are *directly* attached to the scaffold, but those for templates **3**, **6**, and **8** are *not*; at least one vector in the latter group of structures is one bond removed from the core. Thus the highlighted bonds in **3**, **6**, and **8** are closer to *C*β and *C*γ atoms in a side-chain than to *C*α and *C*β's hence they are less constrained than those in the other mimics (illustrated for the oligobenzamide scaffold in Figure 6.5). A consequence of assigning *C*α and *C*β bonds in the original way for **3**, **6**, and **8** is that this covers more

conformational space than, for instance, structure **3'** that focuses on bonds *adjacent* the scaffold. This issue is accentuated for **3** and **6** because all three side-chains are of this type, whereas only one side-chain is impacted for **8** (*ie* the one involving the exocyclic amine).

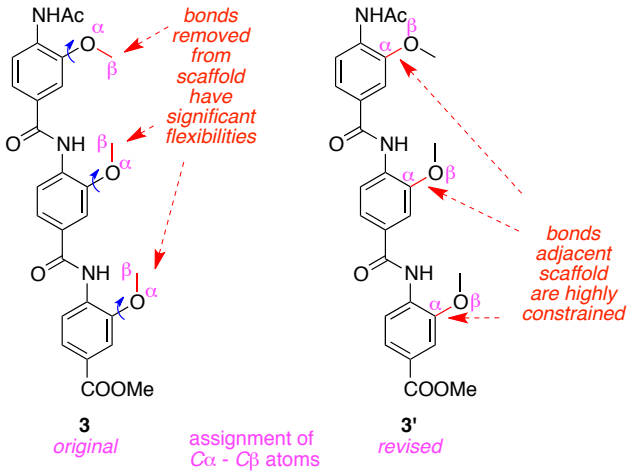


Figure 6.5. Original and revised *C*α and *C*β assignments illustrated for mimic **3**.

As a result of the considerations above, conformational and matching analyses were repeated using structures **3'**, **6'**, and **8'** with *C*α and *C*β assignments revised so that they correspond to vectors attached to the scaffold (Table 6.3). Under these conditions, all the mimics fit somewhat less well on most secondary structures, but conformers of the oligobenzamide system **3'** gave an excellent fit on the ideal α-helix. In fact, that was the best fit identified in this work for any mimic on any secondary structure.

Table 6.3. Matching preferred scaffold conformations using revised *C*α and *C*β assignments[a].

	3_{10}-helix	α-helix	π-helix	β-strand	β-sheet (parallel)	β-sheet (anti-parallel)	sheet-turn-sheet
3'	0.33	0.25	0.36	0.66	0.41	0.39	0.40
6'	0.59	0.66	0.70	-	-	-	0.53
8'	0.99	0.79	0.71	0.88	0.40	0.95	0.40

[a] RMSD ≤ 0.30 Å, red background; 0.31 - 0.50 Å, yellow; 0.51 - 0.70 Å, green.

Figure 5a - d compares overlay data for the ten mimics for each of the three helical secondary structures, with an arbitrary 0.5 Å RMSD cutoff shown by a red line. Thus, for instance, Figure 6.6a shows conformers of **3**, **4**, and **2** overlaid well on an ideal 3_{10}-helix relative to the other mimics; scaffold **3** gave the best overlay and this is shown on the right. Only the oligobenzamide system **3** overlaid with an RMSD of < 0.5 Å on the ideal α-helix (6.6b), and the matching was improved when bonds adjacent the scaffold were considered (**3'**, 6.6c). Similarly, mimic **3'** (and **3**, overlay not shown but data in Table 6.2) gave the best overlays for the π-helix.

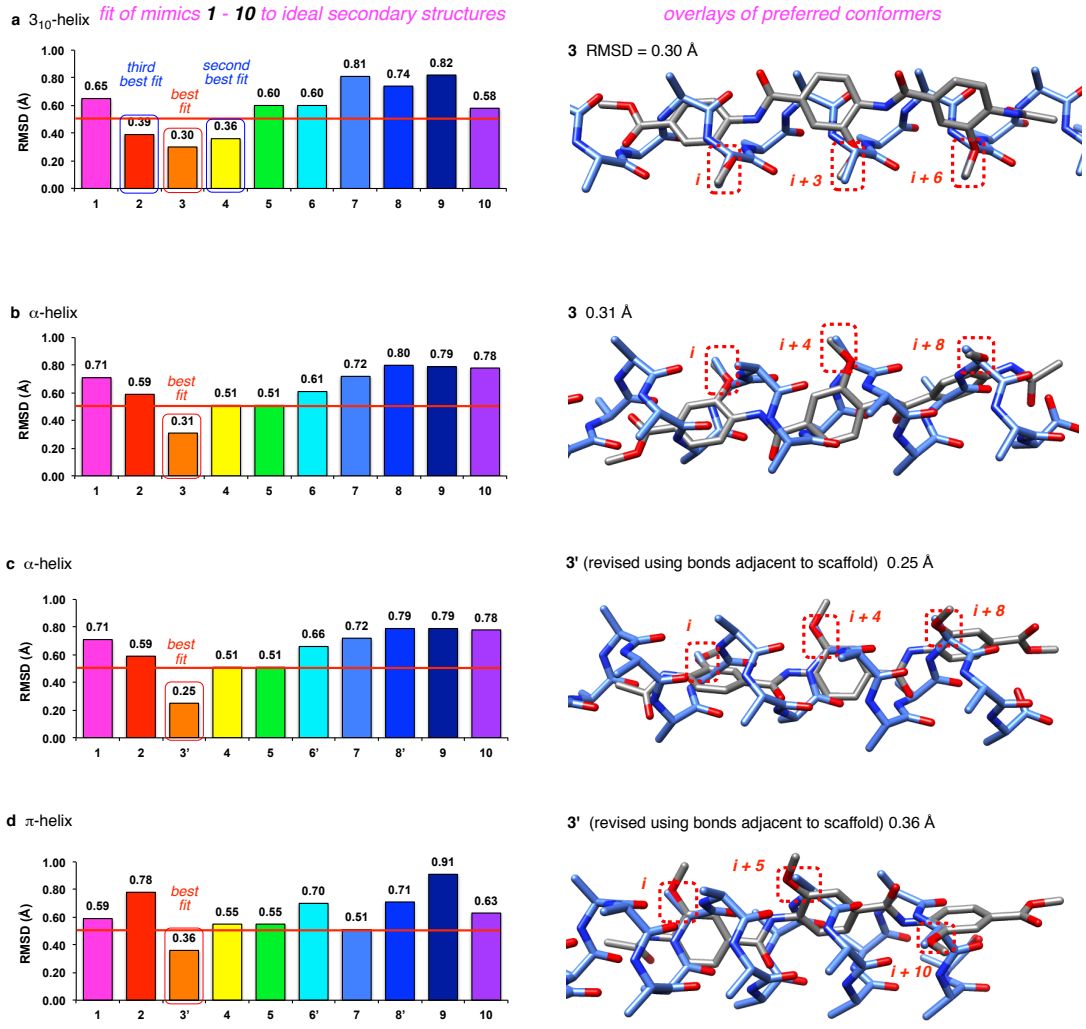


Figure 6.6. Overlays of mimics **1** – **10** on ideal helical structures. The best match for each helical structure is shown on the right.

Figure 6.7 shows overlay data for the mimics on extended, sheet-like structures. There were no good (<0.5 Å RMSD) correspondences for the ideal β-strand (6.7a). As expected, Smith's sheet mimic **9** and the pyrrolinones-pyrrolidine oligomer **10** both gave preferred conformers that matched parallel β-sheets almost equally well (6.7b). However, contrary to expectations: (i) the putative helical mimics **3'** and **8'** also match parallel β-sheets; (ii) scaffolds **9** and **10** do *not* overlay exceptionally well on anti-parallel β-sheets; and, (iii) the oligobenzamide mimic (analyzed using the **3'** or the **3** designation) populates conformers that *do* correspond to anti-parallel β-sheets.

Finally, mimic **7** was shown to have preferred conformers that overlay better on the sheet-turn-sheet than any other mimic; to do this it spans the two sheets and part of the turn region (6.7d). Scaffolds **3**, **3'**, **8'**, **9**, and **10** also had conformers that overlaid well on the sheet-turn-sheet motif.

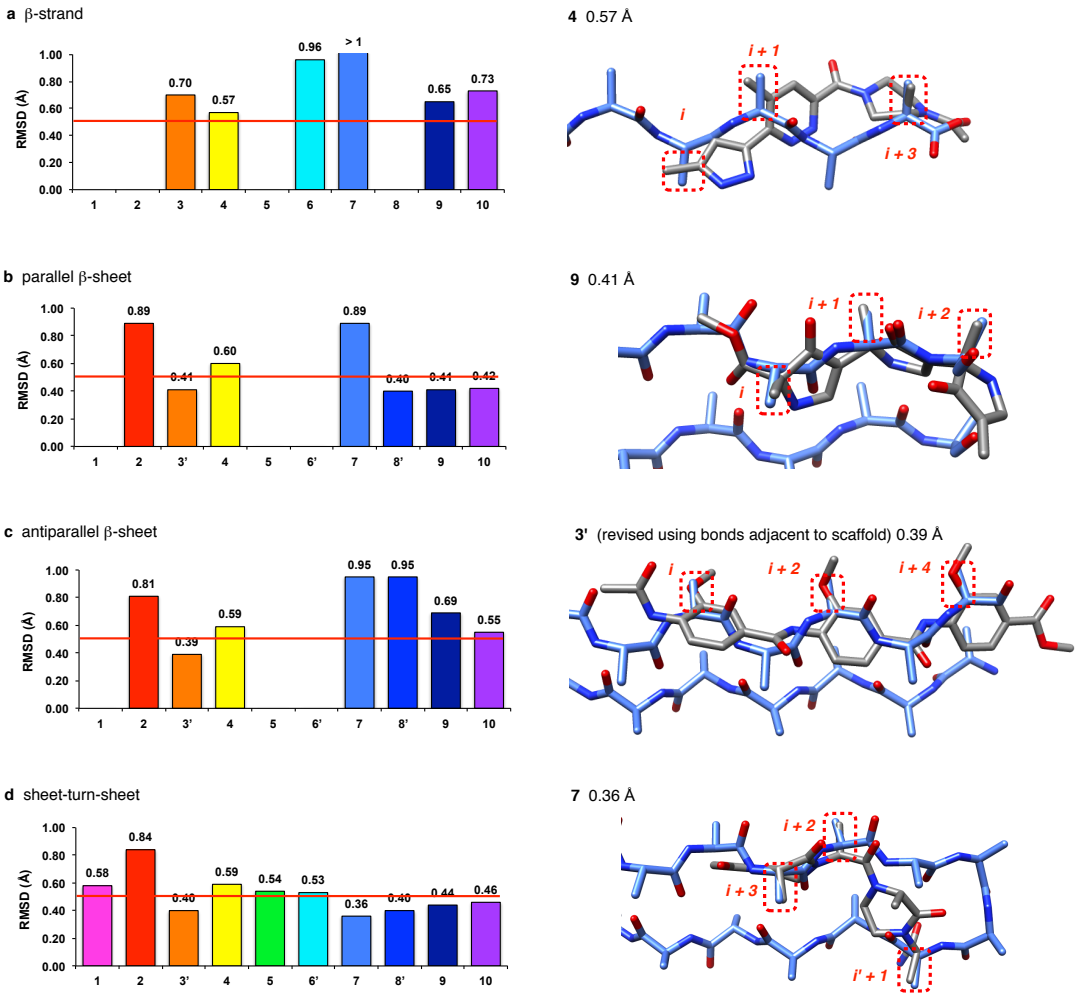


Figure 6.7. Overlays of mimics **1** – **10** on ideal extended sheet-like structures. The best match for each motif is shown on the right.

The strategy used for comparing preferred conformations of the mimics with secondary structures described above does *not* consider the number of conformers that matched well (below RMSD 0.5 Å), or how much more energetic these matching conformers are relative to the lowest energy conformer detected (ΔE). To delve into these issues it is necessary to plot RMSDs *vs* ΔEs for each conformer and Figure 6.8 shows illustrative data.

Figure 6.8 shows the distribution of conformers of the **3'** oligobenzamide description that overlay with 3_{10}-, α-, and π-helices at less than 1.0 Å RMSD and within 3 kcal•mol^{-1} of the lowest energy form observed. All of the 3_{10}-helical conformers that did so with an RMSD of <0.5 Å had an *i*, *i*+3, *i*+6 side-chain correspondence (red dots boxed in red). Smith's analog **9** overlaid on a parallel β-sheet is shown for comparison. This scaffold is the most rigid in the series **1** – **10**, hence it is unsurprising that its conformers cluster tightly; in fact, they aggregate in approximately two groups, each with similar RMSDs. One of these conformational groups overlays parallel β-sheets well with a high population of low energy conformers.

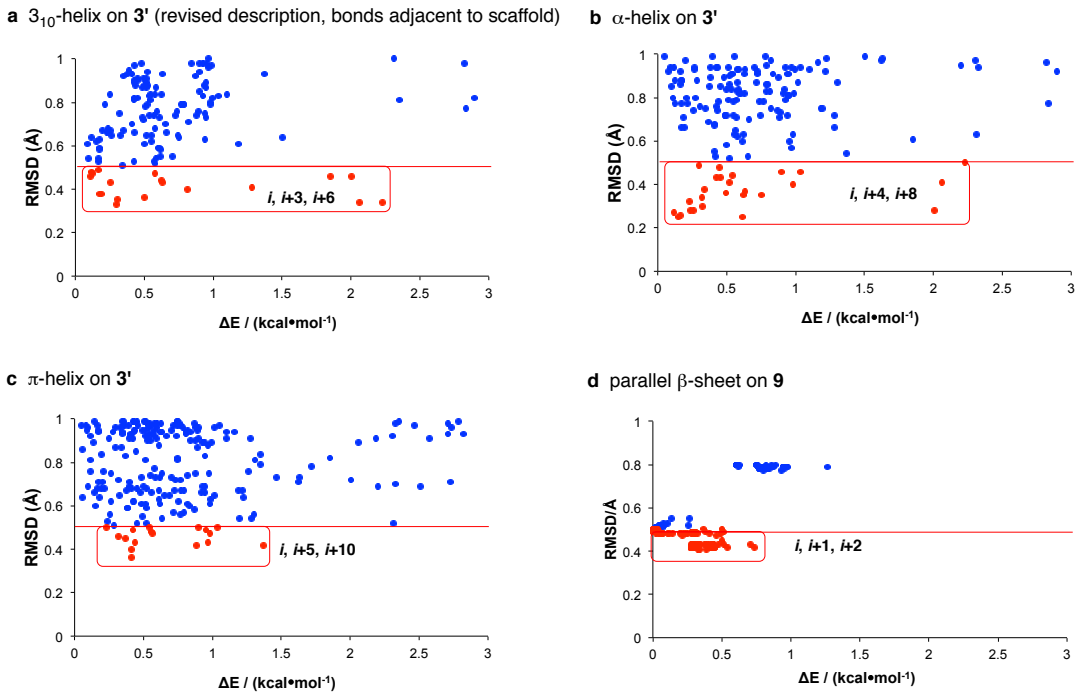


Figure 6.8. Scatter plots of RMSD values of conformers *vs* energies relative to the lowest energy conformer detected (ΔE) indicate how well each mimic populates the featured overlay states. The following numbers of conformers match the indicated secondary structures: **a** compound **3'** on 3_{10}-helix, 19 of 567 conformers, 3.4 %; **b** compound **3'** on α-helix, 26/567, 4.6 %; **c** compound **3'** on π-helix, 17/567, 3 %; compound **9** on β-sheet, 386/600, 64.3 %.

For brevity, this data is summarized in Table 6.4, which gives a semi-quantitative summary of all the dot-plot data for each scaffold type. For instance, no conformers were found for mimic **1** that were below 0.5 Å for overlay on any secondary structure (open circles for whole row), while mimic **10** had some conformers between 1.0 and 3.0 kcal•mol^{-1} that overlaid with β-sheets or strand-turn-strand motifs (purple ticks). Conformers less than 1 kcal•mol^{-1} above the minimum energy one that also matched ideal secondary structures with RMSDs of <0.5 Å were found for mimics **2**, **3** (and **3'**), **4**, **8**, and **9**.

Table 6.4. Summary of data from RMSD/ΔE scatter plots.

	3_{10}-helix	α-helix	π-helix	β-strand	β-sheet (parallel)	β-sheet (anti-parallel)	strand-turn-strand
1	O	O	O	O	O	O	O
2	✓✓	O	O	O	O	O	O
3	✓✓	✓✓	✓✓	O	O	✓✓	✓✓
3'	✓✓	✓✓	✓✓	O	✓✓	✓✓	✓✓
4	✓✓	O	O	O	O	O	O
5	O	O	O	O	O	O	O
6	O	O	O	O	O	O	O
6'	O	O	O	O	O	O	O
7	O	O	O	O	O	O	✓
8	O	O	O	O	O	O	O
8'	O	O	O	O	✓✓	O	✓✓
9	O	O	O	O	✓✓	O	✓✓
10	O	O	O	O	✓	O	✓

✓✓, RMSD < 0.50 Å and ΔE < 1.0 kcal/mol; ✓, RMSD < 0.50 Å and 1.0 kcal/mol < ΔE < 3.0 kcal/mol; O, RMSD > 0.50 Å

6.5 Combinations of Side-chains that Best Fit α-Helical Structures

Scaffolds **1** – **8** were originally designed to resemble α-helices. Overlays of each were shown by their discoverers to illustrate combinations of amino acid side-chains in the secondary structures that were presumed to match with preferred conformations of the mimics. These original assignments of side-chain combinations are shown in blue below each structure in Figure 6.1. Shown in red are the side-chain combinations that correspond to the best overlay on an ideal α-helix as determined using our analysis. *Those assignments of side-chain combinations are different to the ones originally made for all the scaffolds shown*, except for Smith's β-sheet mimic. For our mimic **10** the procedure reported (by us) was the same as that used here,[62] so there is no difference.

Higher energy conformers were probed to determine if the side-chain combinations originally proposed for α-helix mimicry were present in our analyses at higher energies than the one with lowest RMSD. Data for this analysis are shown in Table 6.5.

Red combinations in Table 6.5 are those for the conformer with lowest RMSD overlay on an α-helix. Those shown in purple indicate favored conformations from our analyses that corresponded to side-chain combinations that are *different* to the most favorable conformation and necessarily have higher RMSDs. For instance, our analysis of mimic **1** found the overlay with the lowest RMSD corresponded to matching the *i*, *i*+2, *i*+5 α-helix side-chains. However, some other conformations of **1** matched three other side-chain sets with higher RMSD (*i*, *i*+3, *i*+5; *i*, *i*+4, *i*+5; *i*, *i*-3, *i*-5, the latter being an anti-parallel overlay on the secondary structure, *ie* one that opposes the *N*-to-*C*

polarity). Absence of any side-chain combinations shown in blue for **1** indicates that *none* of the preferred conformers in our analysis corresponded to the amino acid side-chain combinations originally proposed (in this particular case that was *i*, *i*+4, *i*+7).

Overview of Table 6.5 shows *only one case* where a conformer was found that corresponded to the original assignment; specifically, one with an RMSD of 0.91 Å corresponding to the *i*, *i*+3, *i*+7 set originally proposed for mimic **4**. For comparison, Smith's sheet scaffold **9** was included and compared with a parallel β-sheet; the proposed side-chain combinations were identical to the exclusive one found via our analysis (thus could be shown in red or blue and were arbitrarily shown in blue).

Table 6.5. Side-chain combinations for preferred conformer with lowest RMSD relative to the α-helix.

	sequence correspondence for α-helices[a]
1	i, i+2, i+5; i, i+3, i+5; i, i+4, i+5; i, i-3, i-5
2	i, i+2, i+3
3	i, i+4, i+8; i, i+4, i+9; i, i+5, i+8; i, i-4, i-8
3'	i, i+4, i+8
4	i, i+3, i+5; i, i+2, i+5; i, i+4, i+6; i, i+2, i+6; i, i-1, i-5; i, i+3, i+6; i, i+3, i+7 (0.91);
5	i, i+2, i+5; i, i+1, i+5; i, i-4, i-5; i, i+3, i+5; i, i+4, i+5; i, i-3, i-5
6	i, i+2, i+6; i, i-2, i-6
6'	i, i-1, i-5; i, i+1, i+5
7	i, i+1, i+3; i, i+1, i+5; i, i-1, i-3; i, i-1, i+3; i, i+3, i+5; i, i-3, i-6; i, i+1, i-3; i, i-1, i+2; i, i-1, i-5; i, i+2, i+4; i, i+1, i+4
8	i, i+4, i+5; i, i-4, i-5; i, i+2, i+5; i, i-3, i-5; i, i+3, i+5
8'	i, i+1, i+2; i, i-1, i-2
9	i, i+1, i+2 for parallel β-sheet
10	i, i+4, i+8; i, i+1, i+5; i, i+2, i+6; i, i+3, i+7; i, i+4, i+7

[a] red indicates amino acid side-chain combinations for the best matching conformer found in this work; purple indicates other conformers found in this work; blue indicates amino acid side-chain combinations indicated for the original publications. Designations with negatives, *eg* *i*, *i*-3, *i*-5, indicate conformations that overlay antiparallel to the *N*-to-*C* orientation in the secondary structure.

6.6 Helical Mimics Do Not Need to Be Ideal to Perturb Protein-protein Interactions

In several cases, helical mimics featured in this study have been made and tested for their ability to perturb PPIs that involve a helix at an interface. Specifically, the PPIs shown in Table 6 have been assayed using mimics **1**,[41] **3**,[226,233] and **8** (p53/MDM2),[219] **1** (smMLCK/calmodulin),[208] **1**[209,210] and **3**[226] (BakBH3/Bcl-x_L), and **1**[234] (gp41, *C*-/*N*-helical region). Just as this work was submitted examples featuring mimic **3** to inhibit the androgen receptor-coregulator interaction were reported too.[235] Red-shaded regions in Table 6 indicate overlays with low RMSD values while yellow- and green- colored ones, respectively, fitted less well.

Table 6.6 also indicates side-chain correspondences colored to show how they correspond to those predicted here for *ideal* secondary structures, and with the predictions in the original papers. Thus in several cases the side-chain correspondences observed when all conformers of the mimics were overlaid on the PPI helical motifs did *not* match either the predictions based on ideal secondary structures (Table 6.2) or the ones made in the original work; these are shown in black in Table 6.6. However, in the majority of cases (shown in red) the side-chain correlations did match those predicted here for ideal secondary structures. Only in one case, shown in blue for mimic **3** overlaid on p53/MDM2, the original prediction coincided with the side-chain correlations found by our overlay procedure applied to the interface helical motifs.

Most of the data shown in Table 6.6 corresponds to mimic/PPI combinations that have *not* been assayed so far. In two cases a mimic/PPI interface combination seems mutually well suited. In the first, mimic **3** is matched well with BakBH3/Bcl-$x_{L,}$[236]

gp41, and smMLCK/calmodulin, but tests have only been reported for the first. Second, and surprisingly, Smith's system **9** shows an excellent match for the BakBH3/Bcl-x_L; no assays have been reported for this since it has, until now, been regarded as exclusively a β-sheet mimic. Closer examination reveals this overlay is on a helical-terminus where the conformation begins to unwind.

Data in Table 6.6 do *not* mean that the authors of the original papers should have chosen different side-chain correspondences, for various reasons. For instance, some of the side-chains that overlay well when our conformational search routine is used point away from the interface. Thus Table 6.6 should only be used to select mimics to disrupt those PPIs after the orientation of the mimic side-chains have been checked to see if they are appropriate to interact with the protein-binding partner.

Differences between the data in Tables 6.2 (RMSDs for overlays on *ideal* secondary structures) and 6.6 (on *actual* PPI interfaces) reflect the fact that secondary structures in proteins are not ideal. For instance, Table 6.2 indicates mimics **7** – **10** are relatively poor helical mimics, but they overlay on the p53/MDM2 interface with RMSDs that are superior to nearly every entry in Table 6.2 for matching all the mimics on any ideal secondary structure.

Table 6.6. RMSD values for preferred mimic conformations overlaid on PPI components[a].

PPI	1	2	3	3'	4	5	6	6'	7	8	8'	9	10
p53/MDM2[b]	0.57	0.66	0.46	0.55	0.51	0.54	0.50	0.59	0.34	0.43	0.33	0.34	0.47
	i,i+2 ,i+3	i,i+2 ,i+3	i,i+4 ,i+7	i,i+4 ,i+7	i,i+3 ,i+6	i,i+1 ,i+3	i,i+2 ,i+4	i,i+4 ,i+5	i,i+1 ,i+3	i,i+2 ,i+3	i,i+1 ,i+2	i,i+1 ,i+2	i,i+4 ,i+8
smMLCK/calmodulin[c]	0.68	0.60	0.25	0.21	0.35	0.52	0.52	0.57	0.57	0.65	0.77	0.68	0.68
	i,i+2 ,i+5	i,i+2 ,i+3	i,i+4 ,i+8	i,i+4 ,i+8	i,i+3 ,i+5	i,i+2 ,i+5	i,i+2 ,i+6	i,i+4 ,i+5	i,i+6 ,i+5	i,i+4 ,i+5	i,i+1 ,i+2	i,i+1 ,i+3	i,i+4 ,i+8
BakBH3/Bcl-x_L[d]	0.61	0.67	0.28	0.42	0.45	0.61	0.47	0.59	0.43	0.90	0.59	0.23	0.49
	i,i+3 ,i+5	i,i+2 ,i+3	i,i+3 ,i+7	i,i+4 ,i+8	i,i+3 ,i+6	i,i+1 ,i+4	i,i+2 ,i+5	i,i+4 ,i+5	i,i+1 ,i+3	i,i+4 ,i+5	i,i+1 ,i+2	i,i+1 ,i+2	i,i+4 ,i+8
gp41C-/N-helical region[e]	0.73	0.55	0.28	0.26	0.40	0.46	0.54	0.62	0.58	0.71	0.73	0.60	0.72
	i i+3,i +5	i,i+2 ,i+3	i,i+4 ,i+8	i,i+4 ,i+8	i,i+3 ,i+5	i,i+2 ,i+5	i,i+2 ,i+6	i,i+4 ,i+5	i,i+1 ,i+3	i,i+4 ,i+5	i,i+1 ,i+2	i,i+1 ,i+2	i,i+4 ,i+8
predicted from ideal secondary structures	i,i+2 ,i+5	i,i+2 ,i+3	i,i+4 ,i+8	i,i+4 ,i+8	i,i+3 ,i+5	i,i+2 ,i+5	i,i+2 ,i+6	i,i+4 ,i+5	i,i+1 ,i+3	i,i+4 ,i+5	i,i+1 ,i+2	i,i+1 ,i+2	i,i+4 ,i+8
predicted originally	i,i+4 ,i+7	i,i+3 ,i+4	i,i+4 ,i+7	-	i,i+3 ,i+7	i,i+3 ,i+7	i,i+2 ,i+7	-	i,i+4 ,i+7	i,i+3 ,i+7	-	i,i+1 ,i+2	i,i+4 ,i+8

[a] Red background RMSD ≤ 0.30 Å; 0.30 Å < RMSD ≤ 0.50 Å, yellow; 0.50 Å < RMSD ≤ 0.70 Å, green. Side-chain correspondences shown in black did not correspond to the original predictions or those from matching ideal secondary structures as described here, blue denotes correspondence to the original prediction, and red indicates correspondence to the matches deduced here. [b] Helical component shown first throughout; PDB identifier, 1YCR. [c]1CDL. [d]1BXL [e]1AIK

Finally, even if a mimic displaces a helical protein fragment at a PPI this does not prove that the small molecule binds in a helical conformation, and it is hard to confirm that it does by most methods. Binding of a helical mimic to a receptor pocket for a protein helix gives only circumstantial evidence. Hamilton et al proved various terphenyls **1** could influence the interaction of calmodulin with the α-helical domain of smooth muscle myosin light-chain kinase,[208] and used the same scaffold to perturb the α-helical binding domain of Bak BH3 interacting with Bcl-x_L (selectively over p53/HDM2).[210] In the former case the compound was used as a *i*, *i*+3 and *i*+7 mimic, but as a *i*, *i*+4 and *i*+7 mimic in the latter case. HSQC experiments with ^{15}N-labeled Bcl-x_L proved the mimic bound in that binding cleft.[210] Such HSQC experiments are a gold standard in the field; they confirm a mimic binds in the targeted region, and that is still rare in the field of minimalist mimics of secondary structures.[237] In Hamilton's studies those experiments proved the terphenyl associates with the hydrophobic cavity where the helical BH3 peptide binds, but the perturbation of protein ^{15}N-chemical shifts on binding does not reveal the conformation of the bound mimic. Moreover, Hamilton's docking studies indicated the terphenyl mimic could occupy the same hydrophobic cleft as Bak in Bcl-x_L but in a slightly different orientation. In Hamilton's work it was not critical to elucidate the exact binding mode since they achieved their objective: to find a small molecule to perturb Bak/Bcl-x_L. However, for helical mimicry in general, the conformation of the small molecule bound to the protein is interesting. This would require crystallography of the complex, and no group has reported such X-ray data for any of the mimics **1** – **8**. That type of crystallographic data can be hard to obtain for

compounds that bind one protein component with low affinity, and for PPIs where the isolated proteins have different structures compared with the PPI complex.

6.7 Conclusions

Computational simulations of rapidly equilibrating conformational states of minimalist mimics in solution reveal information that cannot be obtained via direct spectroscopic measurements. Preferred simulated conformations of methyl-substituted scaffolds in a continuous dielectric of 80 represent the intrinsic bias of that scaffold with methyl groups in the absence of explicit water molecules. In other words, they represent how the scaffold-core is bias to project $C\alpha$ - $C\beta$ vectors as it surrenders water of solvation and begins to interact with a protein-binding partner.

Orientations of $C\alpha$ and $C\beta$ atoms in preferred conformations of minimalist mimics can be related to the same vectors in secondary structures by automated overlay routines. Computational methods like this can be repeated for hundreds of conformers, making it possible to evaluate a whole conformational ensemble in terms of quantitative RMSD outputs; this is faster and more reliable than could ever be achieved by matching two 3D molecules "by eye" based on a 2D representation on a screen, or by docking a few conformations from large ensembles via MD simulation routines. The key innovation in the research described here is to introduce this type of procedure as a means to evaluate the general applicability of the featured compound types in secondary structure and interface mimicry.

There have been so many papers on minimalist helical mimics that readers who do not follow the field might assume the major challenges must have been surmounted, but this study convinced us that the opposite is true. It is difficult to relate the mimics to their preferred side-chain correspondences, and there are many side-chain correspondences which, based on these simulations, cannot be matched by any of the mimics **1** – **10** with an RMSD of less than 0.5 Å. It would be optimal to have at least one minimalist mimic that can attain a good match (arbitrarily this might be 0.5 Å RMSD) for each side-chain correspondence in every secondary structure, but this has not been achieved. Within the constraints of the methods presented here (limitations of the force fields, assumptions regarding the media for simulations) our simulations indicate that the best mimic of an α-helix is **3**, and that happens to correspond to *i*, *i*+4, *i*+8 side-chain orientations (RMSD of 0.31). In fact, none of the other helical mimics cover other side-chain combinations with an RMSD of <0.50. Another challenge in the field of helical mimicry is to improve the suitability of helical mimics for applications as cell permeable chemical probes and pharmaceutical leads. Compounds containing scaffold **3**, for instance, are unlikely to be cell permeable due to the amide *H*-bond donors. There are ample opportunities for refined design of minimalist secondary structure mimics.

Simulated conformational equilibria of the featured scaffolds reveals that many err towards being universal mimics[46] (several secondary structures represented in one conformational ensemble). Thus, the compounds have the potential to be used as α-helical mimics and to resemble other secondary structures; the featured scaffolds might be used in ways that may not have been obvious before.

Some of the mimics **1** – **8** were probably conceived to match non-ideal helices at specific PPI interfaces. Simulations here indicate that many of the featured helical mimics cannot access conformations that overlay well (*eg* <0.5 Å RMSD) with *ideal* secondary structures, but *can* match distorted helices at particular PPI interfaces. In general, research focused on perturbing PPIs requires close consideration of side-chain orientations *at the particular targeted interface*. We feel that this is the direction the field is already moving in: design and synthesis of *interface mimics*, rather than secondary structure mimics. However, it does not matter if the target conformation is an ideal secondary structure or a completely "non-classical" one observed in a PPI, simulations of the type outlined here will be valuable because evaluation of minimalist mimic conformations with PPI target conformations can reveal information that is not conveniently obtained via spectroscopy. This is, of course, especially true for predictive work to evaluate potential interface mimics before they have been prepared (*cf* Table 6.6). Thus simulations such as these are a possible opening steps in a process that should be followed by checking for the absence of unfavorable backbone interactions of the scaffold with the protein receptor, syntheses, binding assays, and determination of the site of binding.

CHAPTER VII

CONCLUSIONS

Overall, this thesis describes design, synthesis, conformational studies and selected applications of several novel minimalist mimic scaffolds. These scaffolds tend to exist in conformations that can mimic the amino acid side-chain orientations of certain protein secondary structures and PPI interface regions. Some of the mimics are shown to be active in disturbing PPI targets. Specifically, the oligomeric mimics are able to mimic extended interface regions while the constrained cyclic peptide scaffolds can be used to target compact areas at PPIs.

As an extension for the oligo-pyrrolidine-pyrrolidinone oligomeric mimic previously studied in Burgess group, I examined a new oligo-piperidine-piperidinone scaffold. In Chapter II, divergent-convergent synthetic route was developed and it is efficient enough to enable the preparation of three-side-chain mimics with a broad range of side-chain combinations and extended oligomers with up to five amino acid side-chains. The use of β-amino acids and piperidine linker solved the problem of epimerization identified in the synthesis of the oligo-pyrrolidine-pyrrolidinone scaffold. Conformational studies of the new scaffold with X-ray crystallography revealed that the new mimic exist in a helical conformation in solid state. The solid-state conformer is shown to mimic the side-chain arrangements of the *i*, *i + 5* and *i + 10* side-chain sequence of an ideal α-helix. Molecular modelings and data mining (QMD and EKOS) in solution revealed that the scaffold can be a multi-faceted secondary structure mimic

with a preference towards α-helix and β-sheet. Results from data mining with EKO further validated the conclusions from EKOS studies on ideal secondary structures. These results lead us to believe that small molecules derived from this scaffold are promising in disturbing real protein-protein interactions.

Another relevant oligo-piperidine-pyrrolidinone scaffold was studied in Chapter III. This mimic was designed by hybridizing the features in oligo-pyrrolidine-pyrrolidinone and oligo-piperidine-piperidinone scaffolds and synthetic accessibility of this scaffold was evaluated. Similar to the oligo-pyrrolidine-pyrrolidinone scaffold, this mimic suffered from significant epimerization at the chiral centers during the synthesis. Although the application of this mimic is limited in disturbing real PPIs due to the epimerization, an interesting synthetic methodology was discovered during the optimization of the synthetic route and it is further applied to the chemoselective synthesis of a broad range of β-enamino esters, thioesters and amides.

In biological studies, derivatives of the oligomeric mimics were shown to disturb the oligomerization of antithrombin in Chapter IV. These derivatives were chosen based on EKO prediction and synthesized with methods described in Chapter II. The effects of the mimics on the oligomerization state of antithrombin were investigated with native gel electrophoresis. This technique allows the separation of antithrombin oligomers with monomers and provides a fast and convenient way to evaluate the biological activity of the compounds. Besides the initial screening, a modified quantitative native gel electrophoresis procedure was applied in the study of antithrombin oligomerization kinetics and confirmed the importance of mimic concentration and amino acid side-

chains on the scaffold. Moreover, enzyme kinetics assay has successfully differentiated the modes of action between mimics. One mimic tended to induce the formation of short antithrombin oligomers while others favored long oligomers. Finally, transmission electron microscope enables the direct visualization of the linear antithrombin oligomers induced by incubation with the small molecules. These biological studies support the prediction of EKO analysis and represent a significant step forward in validating EKO and EKOS approaches.

Intrigued by the fact that turns are involved in many PPIs, a new constrained cyclic peptide scaffold is also explored in Chapter V of this thesis. The design features a 13-membered macrocyclic ring and an anthranilic acid motif that facilitates the cyclization and increases the conformational rigidity of the macrocycle. We were pleased to find that the scaffold is conformational homogeneous with experiments and with molecular modeling. Furthermore, the macrocyclic ring conformation is not significantly affected by the substituents on amino acid side-chains but has a simple and strong correlation with the stereochemistry of the amino acids based on the 2D NMR studies. In EKOS analysis, different stereomers of the scaffold are found to closely mimic several common turn types. This indicates that the application of this scaffold would not be limited to a certain type of turn structure at PPI interfaces. Besides, after cyclization, the mimics become more drug-like with a lower calculated polar surface area, higher predicted cell permeability and better proteolytic stability. These constrained cyclic peptides described in Chapter V are favorable scaffolds for compact

interface regions. The biological activities of the cyclic peptide derivatives are still under evaluation for selected PPI targets.

In Chapter VI, we have compared the existing key helical mimic scaffolds published by other research groups with our EKOS technique. The major discovery in this research is that although these scaffolds are designed to reproduce the side-chain arrangement of an α-helix, they can potentially mimic other secondary structures besides α-helix, such as 3^{10}-helix, π-helix or β-sheet. Moreover, even for α-helix, those scaffolds are inclined to mimic side-chain sequences other than those claimed in the original publications. These results suggest that researchers have underestimated the potential of these scaffolds and EKOS is useful in revealing new opportunities for existing minimalist mimics. Further experimental research would be needed to find out new applications for these molecules in disturbing PPI targets.

Overall, the designed scaffolds in this thesis show promises in mimicking protein secondary structures and crystallized PPI interface regions. These molecules represent suitable scaffolds in the EKO and EKOS analysis. EKO and EKOS has been proven to be successful in several cases but it still remains to be studied how general these computational methods are in finding small molecules for other protein-protein interaction targets. Further research would be necessary to achieve a better understanding of the scope of EKO and EKOS.

REFERENCES

(1) Wells, J. A.; McClendon, C. L. *Nature* **2007**, *450*, 1001.

(2) Gul, S.; Hadian, K. *Expert Opin. Drug. Discov.* **2014**, *9*, 1.

(3) Higueruelo, A. P.; Jubb, H.; Blundell, T. L. *Curr. Opin. Pharmacol.* **2013**, *13*, 1.

(4) Yohannes, D. In *Annu. Rep. Med. Chem.*; Desai, E., Ed.; Elsevier: 2003, p 295.

(5) Costet, P.; Krempf, M.; Cariou, B. *Trends Biochem. Sci.* **2008**, *33*, 426.

(6) Arkin, M. R.; Wells, J. A. *Nat. Rev. Drug Discovery* **2004**, *3*, 301.

(7) Aeluri, M.; Chamakuri, S.; Dasari, B.; Guduru, S. K. R.; Jimmidi, R.; Jogula, S.; Arya, P. *Chem. Rev.* **2014**, *114*, 4640.

(8) Cochran, A. G. *Cur. Opin. Chem. Bio.* **2001**, *5*, 654.

(9) Vassilev, L. T.; Vu, B. T.; Graves, B.; Carvajal, D.; Podlaski, F.; Filipovic, Z.; Kong, N.; Kammlott, U.; Lukacs, C.; Klein, C.; Fotouhi, N.; Liu, E. A. *Science* **2004**, *303*, 844.

(10) Youle, R. J.; Strasser, A. *Nat. Rev. Mol. Cell Biol.* **2008**, *9*, 47.

(11) Fletcher, S.; Hamilton, A. D. *Curr. Top. Med. Chem.* **2007**, *7*, 922.

(12) Berg, T. *Curr. Opin. Drug Discovery Dev.* **2008**, *11*, 666.

(13) Goldstein, N. I.; Prewett, M.; Zuklys, K.; Rockwell, P.; Mendelsohn, J. *Clin Cancer Res* **1995**, *1*, 1311.

(14) Okines, A.; Cunningham, D.; Chau, I. *Nat. Rev. Clin. Oncol.* **2011**, *8*, 492.

(15) Curran, M. A.; Montalvo, W.; Yagita, H.; Allison, J. P. *Proc. Natl. Acad. Sci.* **2010**, *107*, 4275.

(16) Doemling, A.; Holak, T. A. *Angew. Chem., Int. Ed.* **2014**, *53*, 2286.

(17) Dreher, M. R.; Liu, W.; Michelich, C. R.; Dewhirst, M. W.; Yuan, F.; Chilkoti, A. *J. Natl. Cancer Inst.* **2006**, *98*, 335.

(18) Cabral, H.; Matsumoto, Y.; Mizuno, K.; Chen, Q.; Murakami, M.; Kimura, M.; Terada, Y.; Kano, M. R.; Miyazono, K.; Uesaka, M.; Nishiyama, N.; Kataoka, K. *Nat. Nanotechnol.* **2011**, *6*, 815.

(19) Jain, R. K.; Stylianopoulos, T. *Nat. Rev. Clin. Oncol.* **2010**, *7*, 653.

(20) Rabanel, J. M.; Aoun, V.; Elkin, I.; Mokhtar, M.; Hildgen, P. *Curr. Med. Chem.* **2012**, *19*, 3070.

(21) Dennis, M. S.; Jin, H.; Dugger, D.; Yang, R.; McFarland, L.; Ogasawara, A.; Williams, S.; Cole, M. J.; Ross, S.; Schwall, R. *Cancer Res.* **2007**, *67*, 254.

(22) Saga, T.; Neumann, R. D.; Heya, T.; Sato, J.; Kinuya, S.; Le, N.; Paik, C. H.; Weinstein, J. N. *Proc. Natl. Acad. Sci.* **1995**, *92*, 8999.

(23) Adams, G. P.; Schier, R.; McCall, A. M.; Simmons, H. H.; Horak, E. M.; Alpaugh, R. K.; Marks, J. D.; Weiner, L. M. *Cancer Res.* **2001**, *61*, 4750.

(24) Rudnick, S. I.; Lou, J.; Shaller, C. C.; Tang, Y.; Klein-Szanto, A. J. P.; Weiner, L. M.; Marks, J. D.; Adams, G. P. *Cancer Res.* **2011**, *71*, 2250.

(25) Baluk, P.; Morikawa, S.; Haskell, A.; Mancuso, M.; McDonald Donald, M. *Am. J. Pathol.* **2003**, *163*, 1801.

(26) di Tomaso, E.; Capen, D.; Haskell, A.; Hart, J.; Logie, J. J.; Jain, R. K.; McDonald, D. M.; Jones, R.; Munn, L. L. *Cancer Res.* **2005**, *65*, 5740.

(27) O'Connor, R. *Anticancer Res.* **2007**, *27*, 1267.

(28) Borsi, L.; Balza, E.; Bestagno, M.; Castellani, P.; Carnemolla, B.; Biro, A.; Leprini, A.; Sepulveda, J.; Burrone, O.; Neri, D.; Zardi, L. *Int. J. Cancer* **2002**, *102*, 75.

(29) Carrasco-Triguero, M.; Yi, J.-H.; Dere, R.; Qiu, Z. J.; Lei, C.; Li, Y.; Mahood, C.; Wang, B.; Leipold, D.; Poon, K. A.; Kaur, S. *Bioanalysis* **2013**, *5*, 1007.

(30) van Schouwenburg, P. A.; Bartelds, G. M.; Hart, M. H.; Aarden, L.; Wolbink, G. J.; Wouters, D. *J. Immunol. Methods* **2010**, *362*, 82.

(31) Firer, M. A.; Gellerman, G. *J. Hematol. Oncol.* **2012**, *5*, 70.

(32) Li, J.; Chen, F.; Cona Marlein, M.; Feng, Y.; Himmelreich, U.; Oyen, R.; Verbruggen, A.; Ni, Y. *Target Oncol.* **2012**, *7*, 69.

(33) Gordon, E. M.; Kerwin, J. F. J. *Combinatorial Chemistry and Molecular Diversity in Drug Discovery*; Wiley-Liss: New York City, 1998.

(34) Carr, R. A. E.; Congreve, M.; Murray, C. W.; Rees, D. C. *Drug Discov. Today* **2005**, *10*, 987.

(35) Rees, D. C.; Congreve, M.; Murray, C. W.; Carr, R. *Nature* **2004**, *3*, 660.

(36) Erlanson, D. A.; Braisted, A. C.; Raphael, D. R.; Randal, M.; Stroud, R. M.; Gordon, E. M.; Wells, J. A. *Proc. Natl. Acad. Sci.* **2000**, *97*, 9367.

(37) Czarna, A.; Beck, B.; Srivastava, S.; Popowicz Grzegorz, M.; Wolf, S.; Huang, Y.; Bista, M.; Holak Tad, A.; Domling, A. *Angew. Chem., Int. Ed.* **2010**, *49*, 5352.

(38) Huang, Y.; Wolf, S.; Beck, B.; Kohler, L.-M.; Khoury, K.; Popowicz, G. M.; Goda, S. K.; Subklewe, M.; Twarda, A.; Holak, T. A.; Domling, A. *ACS Chem. Biol.* **2014**, *9*, 802.

(39) Czyzewski, A. M.; Barron, A. E. *AIChE J.* **2008**, *54*, 2.

(40) Fletcher, S.; Hamilton, A. D. *J. Royal. Soc., Interface* **2006**, *3*, 215.

(41) Yin, H.; Lee, G.-i.; Park, H. S.; Payne, G. A.; Rodriguez, J. M.; Sebti, S. M.; Hamilton, A. D. *Angew. Chem. Int. Ed.* **2005**, *44*, 2704.

(42) Yin, H.; Hamilton, A. D. *Angew. Chem., Int. Ed.* **2005**, *44*, 4130.

(43) Yin, H.; Hamilton, A. D. In *Chemical Biology: From Small Molecules to System Biology and Drug Design*; Schreiber, S. L., Kapoor, T. M., Wess, G., Eds.; Wiley-VCH: 2007; Vol. 1, p 250.

(44) Yin, H.; Lee, G.-I.; Hamilton, A. D. *Drug Discovery Res.* **2007**, 281.

(45) Tosovska, P.; Arora, P. S. *Org. Lett.* **2010**, *12*, 1588.

(46) Ko, E.; Liu, J.; Perez, L. M.; Lu, G.; Schaefer, A.; Burgess, K. *J. Am. Chem. Soc.* **2011**, *133*, 462.

(47) Ko, E.; Liu, J.; Burgess, K. *Chem. Soc. Rev.* **2011**, *40*, 4411.

(48) Conte, L. L.; Chothia, C.; Janin, J. *J. Mol. Biol.* **1999**, *285*, 2177.

(49) Keskin, O.; Gursoy, A.; Ma, B.; Nussinov, R. *Chem. Rev.* **2008**, *108*, 1225.

(50) Lao, B. B.; Grishagin, I.; Mesallati, H.; Brewer, T. F.; Olenyuk, B. Z.; Arora, P. S. *Proc. Natl. Acad. Sci.* **2014**, *111*, 7531.

(51) Lao, B. B.; Drew, K.; Guarracino, D. A.; Brewer, T. F.; Heindel, D. W.; Bonneau, R.; Arora, P. S. *J. Am. Chem. Soc.* **2014**, *136*, 7877.

(52) Oh, M.; Lee, J. H.; Wang, W.; Lee, H. S.; Lee, W. S.; Burlak, C.; Im, W.; Hoang, Q. Q.; Lim, H.-S. *Proc. Natl. Acad. Sci.* **2014**, *111*, 11007.

(53) Jung, K.-Y.; Vanommeslaeghe, K.; Lanning Maryanna, E.; Yap Jeremy, L.; Gordon, C.; Wilder Paul, T.; Mackerell Alexander, D., Jr.; Fletcher, S. *Org. Lett.* **2013**, *15*, 3234.

(54) German, E. A.; Ross, J. E.; Knipe, P. C.; Don, M. F.; Thompson, S.; Hamilton, A. D. *Angew. Chem., Int. Ed.* **2015**, *54*, 2649.

(55) Sutherell, C. L.; Thompson, S.; Scott, R. T. W.; Hamilton, A. D. *Chem. Commun.* **2012**, *48*, 9834.

(56) Wyrembak, P. N.; Hamilton, A. D. *J. Am. Chem. Soc.* **2009**, *131*, 4566.

(57) Xin, D.; Ko, E.; Perez, L. M.; Ioerger, T. R.; Burgess, K. *Org. Biomol. Chem.* **2013**, *11*, 7789.

(58) Ko, E.; Raghuraman, A.; Perez, L. M.; Ioerger, T. R.; Burgess, K. *J. Am. Chem. Soc.* **2013**, *135*, 167.

(59) Eguchi, M.; Kahn, M. *Mini-Rev. Med. Chem.* **2002**, *2*, 447.

(60) Angelo, N. G.; Arora, P. S. *J. Am. Chem Soc.* **2005**, *127*, 17134.

(61) Angelo, N. G.; Arora, P. S. *J. Org. Chem.* **2007**, *72*, 7963.

(62) Raghuraman, A.; Ko, E.; Perez, L. M.; Ioerger, T. R.; Burgess, K. *J. Am. Chem. Soc.* **2011**, *133*, 12350.

(63) Raghuraman, A.; Xin, D.; Perez, L. M.; Burgess, K. *J. Org. Chem.* **2013**, *78*, 4823.

(64) Xin, D. Y.; Burgess, K. *Abstr Pap Am Chem S* **2014**, *247*.

(65) Xin, D.; Perez, L. M.; Ioerger, T. R.; Burgess, K. *Angew. Chem. Int. Ed.* **2014**, *53*, 3594.

(66) Xin, D.; Raghuraman, A.; Burgess, K. *J. Org. Chem.* **2015**, *80*, 4450.

(67) Rodriguguez, M. L., M.; Doulut S.; Heitz A.; Martinez, J. *Tetrahedron Letters* **1991**, *32*, 923.

(68) Mosa, F.; Thirsk, C.; Vaultier, M.; Maw, G.; Whiting, A. *Org. Synth.* **2008**, *85*, 219.

(69) Guichard, G.; Chaloin, O.; Cabart, F.; Marin, J.; Zhang, H.; Mihara, H. *Org. Synth.* **2008**, *85*, 147.

(70) Cooper, C. S.; Klock, P. L.; Chu, D. T. W.; Hardy, D. J.; Swanson, R. N.; Plattner, J. J. *J. Med. Chem.* **1992**, *35*, 1392.

(71) Wang, Z.; Miller, E. J.; Scalia, S. J. *Org. Lett.* **2011**, *13*, 6540.

(72) Schobert, R. *Organic Syntheses* **2005**, *82*, No pp given.

(73) Semetey, V.; Moustakas, D.; Whitesides George, M. *Angew. Chem. Int. Ed.* **2006**, *45*, 588.

(74) Driver, R. W.; Hoang, H. N.; Abbenante, G.; Fairlie, D. P. *Org. Lett.* **2009**, *11*, 3092.

(75) O'Connor, S. D.; Smith, P. E.; Al-Obeidi, F.; Pettitt, B. M. *J. Med. Chem.* **1992**, *35*, 2870.

(76) Pettitt, B. M.; Matsunaga, T.; Al-Obeidi, F.; Gehrig, C.; Hruby, V. J.; Karplus, M. *Biophys. J.* **1991**, *60*, 1540.

(77) Bestmann, H. J. *Angew. Chem. Int. Ed.* **1977**, *16*, 349.

(78) Schobert, R.; Dietrich, M.; Mullen, G.; Urbina-Gonzalez, J.-M. *Synthesis* **2006**, 3902.

(79) Schobert, R. *Org. Synth.* **2005**, *82*, 140.

(80) Carpino, L. A. *J. Am. Chem. Soc.* **1993**, *115*, 4397.

(81) Subiros-Funosas, R.; Prohens, R.; Barbas, R.; El-Faham, A.; Albericio, F. *Chem.--Eur. J.* **2009**, *15*, 9394.

(82) Xin, D., K. Burgess *Org. Lett.* **2014**, *16*, 2108.

(83) Jia, Z.; Nagano, T.; Li, X.; Chan, A. S. C. *Eur. J. Org. Chem.* **2013**, *2013*, 858.

(84) Neumann, J. J.; Rakshit, S.; Droege, T.; Wuertz, S.; Glorius, F. *Chem. - Eur. J.* **2011**, *17*, 7298.

(85) Nguyen, H. H.; Kurth, M. J. *Org. Lett.* **2013**, *15*, 362.

(86) Kramer, S.; Dooleweerdt, K.; Lindhardt, A. T.; Rottlander, M.; Skrydstrup, T. *Org. Lett.* **2009**, *11*, 4208.

(87) Zheng, Y.; Li, X.; Ren, C.; Zhang-Negrerie, D.; Du, Y.; Zhao, K. *J. Org. Chem.* **2012**, *77*, 10353.

(88) Suri, M.; Jousseaume, T.; Neumann, J. J.; Glorius, F. *Green Chem.* **2012**, *14*, 2193.

(89) Yamamoto, S.-i.; Okamoto, K.; Murakoso, M.; Kuninobu, Y.; Takai, K. *Org. Lett.* **2012**, *14*, 3182.

(90) Noole, A.; Borissova, M.; Lopp, M.; Kanger, T. *J. Org. Chem.* **2011**, *76*, 1538.

(91) Chun, Y. S.; Xuan, Z.; Kim, J. H.; Lee, S.-g. *Org. Lett.* **2013**, *15*, 3162.

(92) Meng, L.; Wu, K.; Liu, C.; Lei, A. *Chem. Commun.* **2013**, *49*, 5853.

(93) Ke, J.; He, C.; Liu, H.; Li, M.; Lei, A. *Chem. Commun.* **2013**, *49*, 7549.

(94) Toh, K. K.; Wang, Y.-F.; Ng, E. P. J.; Chiba, S. *J. Am. Chem. Soc.* **2011**, *133*, 13942.

(95) Zhao, M.; Wang, F.; Li, X. *Org. Lett.* **2012**, *14*, 1412.

(96) Toh, K. K.; Sanjaya, S.; Sahnoun, S.; Chong, S. Y.; Chiba, S. *Org. Lett.* **2012**, *14*, 2290.

(97) Zhang, Z.-H.; Yin, L.; Wang, Y.-M. *Adv. Synth. Catal.* **2006**, *348*, 184.

(98) Sridharan, V.; Avendano, C.; Menendez, J. C. *Synlett* **2007**, 881.

(99) Knorr, R.; Trzeciak, A.; Bannwarth, W.; Gillessen, D. *Tetrahedron Lett.* **1989**, *30*, 1927.

(100) Coste, J.; Le-Nguyen, D.; Castro, B. *Tetrahedron Lett.* **1990**, *31*, 205.

(101) Vitale, P.; Di Nunno, L.; Scilimati, A. *Synthesis* **2010**, 3195.

(102) Roh, E. J.; Keller, J. M.; Olah, Z.; Iadarola, M. J.; Jacobson, K. A. *Bioorg. Med. Chem.* **2008**, *16*, 9349.

(103) Allen, G. R., Jr., p 337.

(104) Ketcha, D. M.; Wilson, L. J.; Portlock, D. E. *Tetrahedron Lett.* **2000**, *41*, 6253.

(105) Patrick, J. B.; Saunders, E. K. *Tetrahedron Lett.* **1979**, 4009.

(106) Suryavanshi, P. A.; Sridharan, V.; Menendez, J. C. *Org. Biomol. Chem.* **2010**, *8*, 3426.

(107) Li, X.; Taechalertpaisarn, J.; Xin, D.; Burgess, K. *Org. Lett.* **2015**, *17*, 632.

(108) Knaupp Anja, S.; Bottomley Stephen, P. *IUBMB Life* **2009**, *61*, 1.

(109) Yamasaki, M.; Li, W.; Johnson, D. J. D.; Huntington, J. A. *Nature* **2008**, *455*, 1255.

(110) Lomas, D. A.; Carrell, R. W. *Nat. Rev. Genet.* **2002**, *3*, 759.

(111) Takehara, S.; Zhang, J.; Yang, X.; Takahashi, N.; Mikami, B.; Onda, M. *J. Mol. Biol.* **2010**, *403*, 751.

(112) Zhou, A.; Carrell, R. W. *J. Mol. Biol.* **2008**, *375*, 36.

(113) Yamasaki, M.; Sendall, T. J.; Harris, L. E.; Lewis, G. M. W.; Huntington, J. A. *J. Biol. Chem.* **2010**, *285*, 30752.

(114) Ekeowa, U. I.; Freeke, J.; Miranda, E.; Gooptu, B.; Bush, M. F.; Perez, J.; Teckman, J.; Robinson, C. V.; Lomas, D. A. *Proc. Natl. Acad. Sci.* **2010**, *107*, 17146.

(115) Dunstone, M. A.; Whisstock, J. C. In *Methods in Enzymology*; Whisstock, J. C., Bird, P. I., Eds.; Elsevier: 2011; Vol. 501, p 63.

(116) Yamasaki, M.; Sendall, T. J.; Pearce, M. C.; Whisstock, J. C.; Huntington, J. A. *EMBO Rep.* **2011**, *12*, 1011.

(117) Huntington, J. A.; Yamasaki, M. In *Methods in Enzymology*; Whisstock, J. C., Bird, P. I., Eds.; Elsevier, Inc.: 2011; Vol. 501, p 379.

(118) Gooptu, B.; Dickens, J. A.; Lomas, D. A. *Trends Mol. Med.* **2014**, *20*, 116.

(119) Devlin, G. L.; Chow, M. K. M.; Howlett, G. J.; Bottomley, S. P. *J. Mol. Biol.* **2002**, *324*, 859.

(120) Dafforn, T. R.; Mahadeva, R.; Elliott, P. R.; Sivasothy, P.; Lomas, D. A. *J. Biol. Chem.* **1999**, *274*, 9548.

(121) Xin, D.; Holzenburg, A.; Burgess, K. *Chem. Sci.* **2014**, *5*, 4914.

(122) McGovern, S. L.; Caselli, E.; Grigorieff, N.; Shoichet, B. K. *J. Med. Chem.* **2002**, *45*, 1712.

(123) James, E. L.; Bottomley, S. P. *Arch. Biochem. Biophys.* **1998**, *356*, 296.

(124) Mahadeva, R.; Chang, W.-S. W.; Dafforn, T. R.; Oakley, D. J.; Foreman, R. C.; Calvin, J.; Wight, D. G. D.; Lomas, D. A. *J. Clin. Invest.* **1999**, *103*, 999.

(125) Gunnarsson, G. T.; Desai, U. R. *J. Med. Chem.* **2002**, *45*, 1233.

(126) Henry, B. L.; Connell, J.; Liang, A.; Krishnasamy, C.; Desai, U. R. *J. Biol. Chem.* **2009**, *284*, 20897.

(127) Holzenburg, A.; Bewley, M. C.; Wilson, F. H.; Nicholson, W. V.; Ford, R. C. *Nature (London)* **1993**, *363*, 470.

(128) Harris, J. R.; Holzenburg, A. *Micron Microsc. Acta* **1989**, *20*, 223.

(129) Holzenburg, A.; Jones, P. C.; Franklin, T.; Pali, T.; Heimburg, T.; Marsh, D.; Findlay, J. B. C.; Finbow, M. E. *Eur. J. Biochem.* **1993**, *213*, 21.

(130) Krishnan, B.; Gierasch, L. M. *Nat. Struct. Mol. Biol.* **2011**, *18*, 222.

(131) Nowick, J. S. *Acc. Chem. Res.* **1999**, *32*, 287.

(132) Khakshoor, O.; Nowick, J. S. *Curr. Opin. Chem. Biol.* **2008**, *12*, 722.

(133) Gupta, S.; Macala, M.; Schafmeister, C. E. *J. Org. Chem.* **2006**, *71*, 8691.

(134) Fan, E.; Yang, J.; Geib, S. J.; Stoner, T. C.; Hopkins, M. D.; Hamilton, A. D. *J. Chem. Soc., Chem. Commun.* **1995**, 1251.

(135) Peczuh, M. W.; Hamilton, A. D. *Chem. Rev.* **2000**, *100*, 2479.

(136) Mallya, M.; Phillips, R. L.; Saldanha, S. A.; Gooptu, B.; Brown, S. C.; Termine, D. J.; Shirvani, A. M.; Wu, Y.; Sifers, R. N.; Abagyan, R.; Lomas, D. A. *J. Med. Chem.* **2007**, *50*, 5357.

(137) Zhou, A.; Stein, P. E.; Huntington, J. A.; Sivasothy, P.; Lomas, D. A.; Carrell, R. W. *J. Mol. Biol.* **2004**, *342*, 931.

(138) Mahadeva, R.; Dafforn, T. R.; Carrell, R. W.; Lomas, D. A. *J. Biol. Chem.* **2002**, *277*, 6771.

(139) Chang, Y.-P.; Mahadeva, R.; Chang, W.-S. W.; Shukla, A.; Dafforn, T. R.; Chu, Y.-H. *Am. J. Respir. Cell Mol. Biol.* **2006**, *35*, 540.

(140) Skinner, R.; Chang, W.-S. W.; Jin, L.; Pei, X.; Huntington, J. A.; Abrahams, J.-P.; Carrell, R. W.; Lomas, D. A. *J. Mol. Biol.* **1998**, *283*, 9.

(141) Fitton, H. L.; Pike, R. N.; Carrell, R. W.; Chang, W. S. W. *Biological Chemistry* **1997**, *378*, 1059.

(142) Dalsgaard, P. W.; Larsen, T. O.; Frydenvang, K.; Christophersen, C. *J. Nat. Prod.* **2004**, *67*, 878.

(143) Kartha, G.; Ambady, G.; Shankar, P. V. *Nature* **1974**, *247*, 204.

(144) Zheng, J.; Xu, Z.; Wang, Y.; Hong, K.; Liu, P.; Zhu, W. *J. Nat. Prod.* **2010**, *73*, 1133.

(145) Wels, B.; Kruijtzer, J. A. W.; Liskamp, R. M. J. *Org. Lett.* **2002**, *4*, 2173.

(146) Takeuchi, Y.; Marshall, G. R. *J. Am. Chem. Soc.* **1998**, *120*, 5363.

(147) El Haddadi, M.; Cavelier, F.; Vives, E.; Azmani, A.; Verducci, J.; Martinez, J. *J. Pept. Sci.* **2000**, *6*, 560.

(148) Cini, E.; Botta, C. B.; Rodriquez, M.; Taddei, M. *Tetrahedron Lett.* **2009**, *50*, 7159.

(149) Ngu-Schwemlein, M.; Zhou, Z.; Bowie, T.; Eden, R. *J. Mol. Struct.* **2003**, *655*, 59.

(150) Meutermans, W. D. F.; Bourne, G. T.; Golding, S. W.; Horton, D. A.; Campitelli, M. R.; Craik, D.; Scanlon, M.; Smythe, M. L. *Organic Lett.* **2003**, *5*, 2711.

(151) Horton, D. A.; Bourne, G. T.; Coughlan, J.; Kaiser, S. M.; Jacobs, C. M.; Jones, A.; Ruehmann, A.; Turner, J. Y.; Smythe, M. L. *Org. Biomol. Chem.* **2008**, *6*, 1386.

(152) Wong, C. T. T.; Lam, H. Y.; Song, T.; Chen, G.; Li, X. *Angew. Chem.* **2013**, *52*, 10212.

(153) Horne, W. S.; Olsen, C. A.; Beierle, J. M.; Montero, A.; Ghadiri, M. R. *Angew. Chem. Int. Ed.* **2009**, *48*, 4718.

(154) Pease, L. G.; Watson, C. *J. Am. Chem. Soc.* **1978**, *100*, 1279.

(155) Bach, A. C., II; Bothner-By, A. A.; Gierasch, L. M. *J. Am. Chem. Soc.* **1982**, *104*, 572.

(156) Bruch, M. D.; Noggle, J. H.; Gierasch, L. M. *J. Am. Chem. Soc.* **1985**, *107*, 1400.

(157) Ehrlich, A.; Heyne, H.-U.; Winter, R.; Beyermann, M.; Haber, H.; Carpino, L. A.; Bienert, M. *J. Org. Chem.* **1996**, *61*, 8831.

(158) Kessler, H.; Kutscher, B. *Tetrahedron Lett.* **1985**, *26*, 177.

(159) Mierke, D. F.; Kurz, M.; Kessler, H. *J. Am. Chem. Soc.* **1994**, *116*, 1042.

(160) Haubner, R.; Gratias, R.; Diefenbach, B.; Goodman, S. L.; Jonczyk, A.; Kessler, H. *J. Am. Chem. Soc.* **1996**, *118*, 7461.

(161) Montero, A.; Beierle, J. M.; Olsen, C. A.; Ghadiri, M. R. *J. Am. Chem. Soc.* **2009**, *131*, 3033.

(162) Olsen, C. A.; Montero, A.; Leman, L. J.; Ghadiri, M. R. *ACS Med. Chem. Lett.* **2012**, *3*, 749.

(163) Beierle, J. M.; Horne, W. S.; van Maarseveen, J. H.; Waser, B.; Reubi, J. C.; Ghadiri, M. R. *Angew. Chem. Int. Ed.* **2009**, *48*, 4725.

(164) Hoang, H. N.; Driver, R. W.; Beyer, R. L.; Malde, A. K.; Le, G. T.; Abbenante, G.; Mark, A. E.; Fairlie, D. P. *Angew. Chem., Int. Ed.* **2011**, *50*, 11107.

(165) Glenn, M.; Kelso, M.; Tyndall, J.; Fairlie, D. *J. Am. Chem. Soc.* **2003**, *125*, 640.

(166) Igarashi, Y.; Hanafusa, T.; Gohda, F.; Peterson, S.; Bills, G. *Mycology* **2014**, *5*, 102.

(167) Hamada, Y.; Nakao, K.; Shioiri, T. *Pept. Chem.* **1988**, 351.

(168) Kobayashi, R.; Samejima, Y.; Nakajima, S.; Kawai, K.; Udagawa, S. *Chem. Pharm. Bull.* **1987**, *35*, 1347.

(169) Schmeda-Hirschmann, G.; Hormazabal, E.; Rodriguez, J. A.; Theoduloz, C. *Z. Naturforsch., C J. Biosci.* **2008**, *63*, 383.

(170) Zhang, Y.-G.; Xia, X.-K.; Yuan, W.-P.; Liu, X.; Zhang, M.-S.; Meng, X.-M.; Wang, X.-J.; Liu, C.-H. *Z. Kristallogr. - New Cryst. Struct.* **2010**, *225*, 236.

(171) Yamazaki, M.; Horie, Y.; Bae, K.; Maebayashi, Y.; Jisai, Y.; Fujimoto, H. *Chem. Pharm. Bull.* **1987**, *35*, 2122.

(172) Convert, O.; Mazaleyrat, J. P.; Wakselman, M.; Morize, I.; Reboud-Ravaux, M. *Biopolymers* **1990**, *30*, 583.

(173) Lewer, P.; Graupner, P. R.; Hahn, D. R.; Karr, L. L.; Duebelbeis, D. O.; Lira, J. M.; Anzeveno, P. B.; Fields, S. C.; Gilbert, J. R.; Pearce, C. *J. Nat. Prod.* **2006**, *69*, 1506.

(174) Zhang, Y.; Liu, S.; Liu, H.; Liu, X.; Che, Y. *J. Nat. Prod.* **2009**, *72*, 1364.

(175) Zhou, L.-N.; Gao, H.-Q.; Cai, S.-X.; Zhu, T.-J.; Gu, Q.-Q.; Li, D.-H. *Helv. Chim. Acta* **2011**, *94*, 1065.

(176) Chen, M.; Shao, C.-L.; Fu, X.-M.; Kong, C.-J.; She, Z.-G.; Wang, C.-Y. *J. Nat. Prod.* **2014**, *77*, 1601.

(177) Mang, C.-Y.; Zhao, Y.; Li, H.-F.; Lan, H.; Yan, Y.; Yang, M.-H. *Mol. Phys.* **2015**, *113*, 104.

(178) Peng, J.; Gao, H.; Zhang, X.; Wang, S.; Wu, C.; Gu, Q.; Guo, P.; Zhu, T.; Li, D. *J. Nat. Prod.* **2014**, *77*, 2218.

(179) Masuda, Y.; Tanaka, R.; Kai, K.; Ganesan, A.; Doi, T. *J. Org. Chem.* **2014**, *79*, 7844.

(180) Mazaleyrat, J. P.; Reboud-Ravaux, M.; Wakselman, M. *Int. J. Pept. Protein Res.* **1987**, *30*, 622.

(181) Boggetto, N.; Vilain, A. C.; Montagne, J. J.; Reboud-Ravaux, M.; Mazaleyrat, J. P.; Xie, J.; Wakselman, M. *Bull. Soc. Chim. Fr.* **1994**, *131*, 152.

(182) Kumar, K. H.; Boja, P.; Belagali, S. L. *Indian J. Heterocycl. Chem.* **2002**, *12*, 149.

(183) Kai, K.; Yoshikawa, H.; Kuo, Y.-H.; Akiyama, K.; Hayashi, H. *Biosci., Biotechnol., Biochem.* **2010**, *74*, 1309.

(184) Zheng, J.; Zhu, H.; Hong, K.; Wang, Y.; Liu, P.; Wang, X.; Peng, X.; Zhu, W. *Org. Lett.* **2009**, *11*, 5262.

(185) Kuo, Y.-H.; Kai, K.; Akiyama, K.; Hayashi, H. *Tetrahedron Lett.* **2012**, *53*, 429.

(186) Mazaleyrat, J. P.; Reboud-Ravaux, M.; Montagne, J. J.; Convert, O.; Wakselman, M. *Colloq. INSERM* **1989**, *174*, 333.

(187) Cerrini, S.; Gavuzzo, E.; Lucente, G.; Pinnen, F. *Int. J. Pept. Protein Res.* **1988**, *31*, 447.

(188) Cerrini, S.; Gavuzzo, E.; Lucente, G.; Pinnen, F.; Zanotti, G. *Int. J. Pept. Protein Res.* **1989**, *34*, 6.

(189) Pinnen, F.; Di Muro, A.; Zanotti, G.; Lucente, G. *Int. J. Pept. Protein Res.* **1987**, *30*, 388.

(190) Feigel, M.; Lugert, G. *Liebigs Ann. Chem.* **1989**, 1089.

(191) Akazome, M.; Enzu, M.; Takagi, K.; Matsumoto, S. *Chirality* **2011**, *23*, 568.

(192) Xu, J.; Zhao, S.; Yang, X. *Nat. Prod. Res.* **2014**, *28*, 994.

(193) He, F.; Bao, J.; Zhang, X.-Y.; Tu, Z.-C.; Shi, Y.-M.; Qi, S.-H. *J. Nat. Prod.* **2013**, *76*, 1182.

(194) Eid, C. N.; Nicas, T. I.; Mullen, D. L.; Loncharich, R. J.; Paschal, J. W. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 2087.

(195) Han, G.; Tamaki, M.; Hruby, V. J. *J. Pept. Res.* **2001**, *58*, 338.

(196) Baxter, N. J.; Williamson, M. P. *J. Biomol. NMR* **1997**, *9*, 359.

(197) Cierpicki, T.; Otlewski, J. *J. Biomol. NMR* **2001**, *21*, 249.

(198) Lingard, H.; Han, J. T.; Thompson, A. L.; Leung, I. K. H.; Scott, R. T. W.; Thompson, S.; Hamilton, A. D. *Angew. Chem., Int. Ed.* **2014**, *53*, 3650.

(199) Steffel, L. R.; Cashman, T. J.; Reutershan, M. H.; Linton, B. R. *J. Am. Chem. Soc.* **2007**, *129*, 12956.

(200) Ludvigsen, S.; Andersen, K. V.; Poulsen, F. M. *J. Mol. Biol.* **1991**, *217*, 731.

(201) Claridge, T. D. W. *High-Resolution NMR Techniques in Organic Chemistry*; Pergamon: Oxford, 1999.

(202) Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. *J. Comput. Chem.* **1990**, *11*, 440.

(203) Jurasek, L.; Johnson, P.; Olafson, R. W.; Smillie, L. B. *Can. J. Biochem.* **1971**, *49*, 1195.

(204) Jorgensen, W. L.; Duffy, E. M. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 1155.

(205) Duffy, E. M.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2000**, *122*, 2878.

(206) Veber, D. F.; Johnson, S. R.; Cheng, H.-Y.; Smith, B. R.; Ward, K. W.; Kopple, K. D. *J. Med. Chem.* **2002**, *45*, 2615.

(207) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Delivery Rev.* **1997**, *23*, 3.

(208) Orner, B. P.; Ernst, J. T.; Hamilton, A. D. *J. Am. Chem. Soc.* **2001**, *123*, 5382.

(209) Kutzki, O.; Park, H. S.; Ernst, J. T.; Orner, B. P.; Yin, H.; Hamilton, A. D. *J. Am. Chem. Soc.* **2002**, *124*, 11838.

(210) Yin, H.; Lee, G.; Sedey, K. A.; Kutzki, O.; Park, H. S.; Orner, B. P.; Ernst, J. T.; Wang, H.-G.; Sebti, S. M.; Hamilton, A. D. *J. Am. Chem. Soc.* **2005**, *127*, 10191.

(211) Kim, I. C.; Hamilton, A. D. *Org. Lett.* **2006**, *8*, 1751.

(212) Ahn, J.-M.; Han, S.-Y. *Tetrahedron Lett.* **2007**, *48*, 3543.

(213) Shaginian, A.; Whitby, L. R.; Hong, S.; Hwang, I.; Farooqi, B.; Searcey, M.; Chen, J.; Vogt, P. K.; Boger, D. L. *J. Am. Chem. Soc.* **2009**, *131*, 5564.

(214) Plante, J.; Campbell, F.; Malkova, B.; Kilner, C.; Warriner, S. L.; Wilson, A. J. *Org. Biomol. Chem.* **2008**, *6*, 138.

(215) Volonterio, A.; Moisan, L.; Rebek, J., Jr. *Org. Lett.* **2007**, *9*, 3733.

(216) Moisan, L.; Odermatt, S.; Gombosuren, N.; Carella, A.; Rebek, J., Jr. *Eur. J. Org. Chem.* **2008**, 1673.

(217) Maity, P.; Koenig, B. *Org. Lett.* **2008**, *10*, 1473.

(218) Marimganti, S.; Cheemala, M. N.; Ahn, J.-M. *Org. Lett.* **2009**, *11*, 4418.

(219) Lee Ji, H.; Zhang, Q.; Jo, S.; Chai Sergio, C.; Oh, M.; Im, W.; Lu, H.; Lim, H.-S. *J Am Chem Soc* **2011**, *133*, 676.

(220) Davis, J. M.; Tsou, L. K.; Hamilton, A. D. *Chem. Soc. Rev.* **2007**, *36*, 326.

(221) Bax, A.; Davis, D. G. *J. Mag. Resonance* **1985**, *63*, 207.

(222) Kessler, H.; Griesinger, C.; Kerssebaum, R.; Wagner, K.; Ernst, R. R. *J. Am. Chem. Soc.* **1987**, *109*, 607.

(223) Prabhakaran, P.; Azzarito, V.; Jacobs, T.; Hardie, M. J.; Kilner, C. A.; Edwards, T. A.; Warriner, S. L.; Wilson, A. J. *Tetrahedron* **2012**, *68*, 4485.

(224) Marimganti, S.; Cheemala, M. N.; Ahn, J.-M. *Org. Lett.* **2009**, *11*, 4418.

(225) Kulikov, O. V.; Hamilton, A. D. *RSC Adv.* **2012**, *2*, 2454.

(226) Azzarito, V.; Prabhakaran, P.; Bartlett, A. I.; Murphy, N. S.; Hardie, M. J.; Kilner, C. A.; Edwards, T. A.; Warrinera, S. L.; Wilson, A. J. *Org. Biomol. Chem.* **2012**, *10*, 6469.

(227) Fuller, J. C.; Jackson, R. M.; Edwards, T. A.; Wilson, A. J.; Shirts, M. R. *PLoS One* **2012**, *7*, e43253.

(228) Yap, J. L.; Cao, X.; Vanommeslaeghe, K.; Jung, K.-Y.; Peddaboina, C.; Wilder, P. T.; Nan, A.; MacKerell, A. D.; Smythe, W. R.; Fletcher, S. *Org. Biomol. Chem.* **2012**, *10*, 2928.

(229) Xin, D. Y.; Ko, E.; Perez, L. M.; Burgess, K. *Abstr Pap Am Chem S* **2013**, *245*.

(230) Smith, A. B., III; Charnley, A. K.; Hirschmann, R. *Acc. Chem. Res.* **2011**, *44*, 180.

(231) Smith, W. W.; Bartlett, P. A. *J. Am. Chem. Soc.* **1998**, *120*, 4622.

(232) Lauri, G.; Bartlett, P. A. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 51.

(233) Plante, J. P.; Burnley, T.; Malkova, B.; Webb, M. E.; Warriner, S. L.; Edwards, T. A.; Wilson, A. J. *Chem. Commun.* **2009**, 5091.

(234) Ernst, J. T.; Kutzki, O.; Debnath, A. K.; Jiang, S.; Lu, H.; Hamilton, A. D. *Angew. Chem. Int. Ed.* **2002**, *41*, 278.

(235) Ravindranathan, P.; Lee, T.-K.; Yang, L.; Centenera, M. M.; Butler, L.; Tilley, W. D.; Hsieh, J.-T.; Ahn, J.-M.; Raj, G. V. *Nat. Commun.* **2013**, *4*, 1.

(236) Ahn, J.-M.; (University of Texas System, USA). Application: US, 2010, p 1.

(237) Campbell, F.; Plante, J. P.; Edwards, T. A.; Warriner, S. L.; Wilson, A. J. *Org. Biomol. Chem.* **2010**, *8*, 2344.

APPENDIX A

GENERAL PROCEDURES

A. General Experimental Procedures

All reactions were carried out under nitrogen atmosphere with dry solvents under anhydrous conditions. Glassware was dried in an oven at 140 ^{o}C for minimum 6 h prior to use. Dry solvents were obtained by passing the previously degassed solvents through activated alumina columns. Reagents were purchased at a high commercial quality (typically 97 % or higher) and used without further purification, unless otherwise stated. Flash chromatography was performed using silica gel (230–600 mesh). Analytical thin layer chromatography (TLC) was carried out on Merck silica gel plates with QF-254 indicator and visualized by UV or potassium permanganate stains. ^{1}H and ^{13}C spectra were recorded on a 300 MHz or a 400 MHz spectrometer and were calibrated using residual non-deuterated solvent as an internal reference. Chemical shifts (δ) are reported in ppm, and coupling constants (J) are given in Hz. The following abbreviations were used to explain the multiplicities: s = singlet, d = doublet, t = triplet, q = quartet, dd = double doublet, dq = double quartet m = multiplet, br = broad. HRMS were obtained using ESI or MALDI ionization. Melting points were recorded on an automated melting point apparatus and are uncorrected.

All the HPLC analyses were carried out with UV detection monitored at 254 nm. Analytical reversed phase HPLC analyses were performed with a 150 x 4.6 mm C-18 column using gradient conditions (10 – 90 % acetonitrile in water, flow rate = 0.75

mL/min). Chiralpak AD (250 x 4.6 mm ID) column was utilized for the chiral HPLC analysis (hexanes : isopropyl alcohol 85 : 15, flow rate = 1 mL / min).

All the UV spectrums were recorded on a UV spectrometer using a 10 mm quartz cuvette at 20 μM in MeOH. Circular Dichroism spectrums were recorded on a CD spectrometer using a 2 mm quartz cuvette at 200 μM in MeOH.

B. General Procedure for Quenched Molecular Dynamics

NAMD[1] was used for the molecular simulations performed in this work. Explicit atom representations were used throughout the study. The protein structure files (PSF) for all the peptidomimetics were built using Discovery Studio 2.5 (Accelrys Inc) using the CHARMm force field.

Quenched molecular dynamics simulations were performed using the CHARMm force field as implemented in Discovery Studio 2.5. All four molecules were modeled as neutral compounds in a dielectric continuum of 80 (simulating H_2O). Thus, the starting conformers were minimized using 3000 steps of conjugate gradient. The minimized structures were then subjected to heating, equilibration, and dynamics simulation. Throughout, the equations of motions were integrated using the Verlet algorithm with a time step 1 fs. Each peptidomimetic was heated to 1000 K over 10 ps and equilibrated for another 10 ps at 1000 K, then molecular dynamics runs were performed for a total time of 600 ps with trajectories saved every 1 ps. The resulting 600 structures were thoroughly minimized using 1000 steps of SD followed by 3000 steps of conjugate

gradient. Structures with energies less than 3.0 kcal mol^{-1} relative to the global minimum were selected for further analysis.

The VMD[2] package was used to display, overlay, and classify the selected structures into conformational groups. The best clustering was obtained using a grouping method based on calculation of RMS deviation of a subset of atoms, in this study these were the Cα - and Cβ- atoms. Thus, threshold cutoff values 0.5 Å were selected to obtain families with reasonable homogeneity. The lowest energy conformation from each family was considered to be a typical representative of the family as a whole.

C. Procedures for Matching on Ideal Secondary Structures and Crystal Structure

Standard template for overlays with 3_{10}-helix, α-helix, π-helix, and β-strand were obtained from Discovery Studio 2.5. Parallel β-sheet, anti-parallel β-sheet and sheet/turn/sheet templates were obtained by modified β-sheet builder (http://www-lbit.iro.umontreal.ca/bBuilder/index.html).

After minimization in the QMD process, the conformers were grouped into families base on their $C\alpha$ - $C\beta$ coordinates. The process of systematically matching preferred conformers with secondary structures was performed in the following way. All the conformers within 3.0 kcal/mol were considered to be "preferred".

Each of these conformers was overlaid on ideal secondary structures, the crystal structure or the NMR structure using an in house generated algorithm that compared

$C\alpha$ - $C\beta$ coordinates of the side chains which generates a list of structures ranked in terms of the RMSD for the overlay process.

D. Procedures for Data Mining of 3D Complex Database

The data mining of 3D complex database was performed for ***LLL*-1aaa** according to the procedure described previously.[3] The overlays were analyzed by DSSP to get the secondary structure information for individual residues.

APPENDIX B

EXPERIMENTAL PROCEDURES FOR CHAPTER II

A. General Procedures

General Procedure for Boc Deprotection:

The substrate was dissolved in 1:1 TFA/dichloromethane to give 0.1 M concentration at 0 ºC and the reaction was stirred for 30 min at room temperature. Toluene (30 mL) was added and the solution was concentrated. Residual TFA was azeotroped 3 times with toluene (3 × 30 mL) and the residue was placed under high vacuum for 3 h to give the crude mixture. The crude mixture was purified with procedures indicated for specific compounds.

General Procedure for Cbz Deprotection:

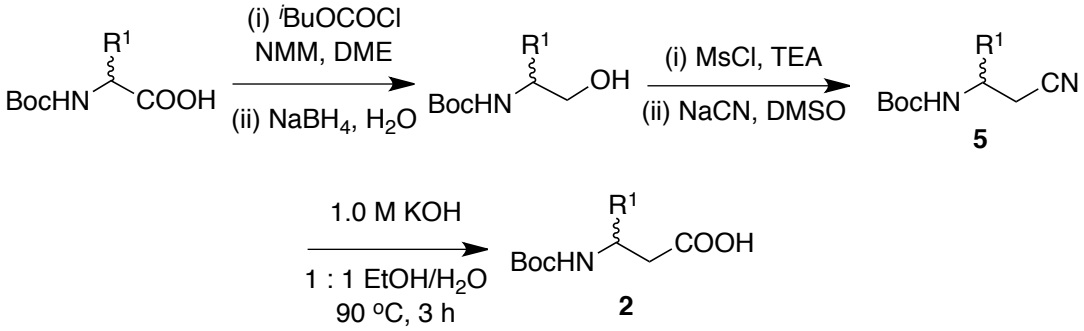
To a solution of substrate in methanol (0.1 M) under nitrogen was added 10 wt % Pd/C (0.1 equiv. Pd). The reaction was placed under an atmosphere of hydrogen (1 atm, balloon) for 12 h. After the reaction finished, the flask was purged with N_2. The reaction mixture was filtered over a celite pad and concentrated to afford the crude product. The crude mixture was purified with procedures indicated for specific compounds.

General Procedure for Coupling Reaction:

To a 0.5 M solution of 1.0 equiv. of the amine in CH_2Cl_2 was added 100.0 equiv. of anhydrous $KHCO_3$ powder. 1.0 equiv. of the vinyl chloride was then added as a 0.5 M

solution in CH_2Cl_2 to the reaction mixture. While stirring at room temperature, the solvent was removed with nitrogen. The resulting solid was heated up to the indicated temperature under N_2 for the indicated time. During the reaction, at about half of the indicated time, 0.5 mL anhydrous THF was added, stirred for 10 mins and then removed with nitrogen. After the reaction finished, CH_2Cl_2 was added and the inorganic solids were removed by filtration. The solution was concentrated to give the crude product.

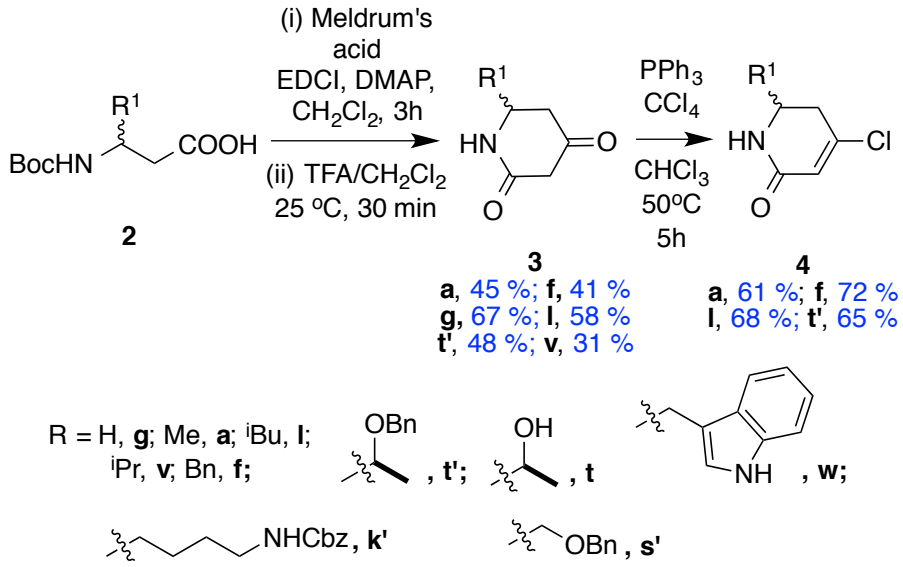
B. Synthesis of Compounds 1



General Procedure for the Syntheses of 5

The *N*-Boc-protected amino alcohol was prepared from *N*-Boc-protected amino acid with literature procedures.[4] The crude *N*-protected amino nitrile compounds were synthesized on large scale from corresponding *N*-Boc-protected amino alcohols with the procedure from the literature.[5] **5a**, **5l**, **5v**, **5g'**, **5f** can be obtained by crystallization of the crude mixture from CH_2Cl_2/hexanes. **5s'**, **5t'**, **5w**, **5k'** can be purified by washing the ether solution of crude product 3 times with saturated $NaHCO_3$ solution and 1 time with

water before drying over $MgSO_4$ and the removal of solvent. In some cases, *D*-amino acid derivatives were used as the starting material.



General Procedure for the Syntheses of 2

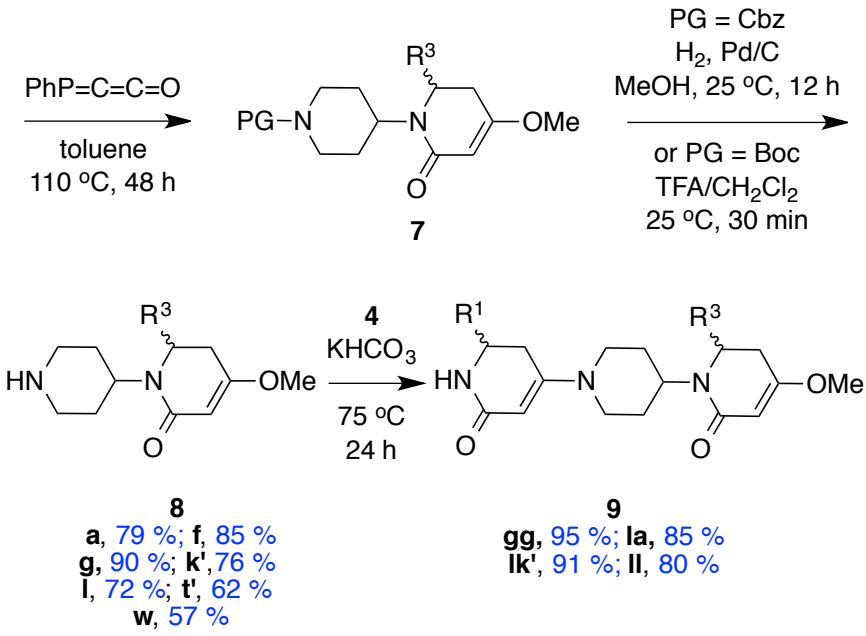
5 (20 mmol) was dissolved in 20 mL EtOH at room temperature and 20 mL 2 M KOH aqueous solution was added in one portion. The resulting mixture was stirred at 90 °C until the reaction was complete (usually 3 h was enough). The mixture was cooled down to room temperature and EtOH was removed under vacuum. The crude product was cooled to 0 °C and 1 M $KHSO_4$ aqueous solution was slowly added until the pH of the solution reached about 2-3. The product was extracted with Et_2O three times (80 mL × 3). The combined Et_2O solution was washed twice with water (100 mL × 2) and brine (100 mL). The ether phase was dried over $MgSO_4$ and concentrated under vacuum. The residue was used directly in the synthesis of **3** without further purification.

General Procedure for the Syntheses of 3[6-8]

To a stirred solution of meldrum's acid (476 mg, 3.3 mmol, 1.1 equiv) and DMAP (550 mg, 4.5 mmol, 1.5 equiv) at 0 ºC in dichloromethane (15 mL) was added N-Boc-β-amino acid **2** (3.0 mmol, 1.0 equiv) in one portion. EDCI (978 mg, 5.1 mmol, 1.7 equiv) was added in one portion and the reaction mixture was stirred at 25 ºC for 5 h. The yellow reaction mixture was transferred to a separatory funnel and diluted with dichloromethane (80 mL) and washed with cold 5 % $KHSO_4$ (3 × 75 mL) and brine (75 mL). The organic layer was dried over $MgSO_4$ and filtered. The filtrate was concentrated under vacuum and 60 mL ethyl acetate was added. After refluxing for 5 h under N_2, the solution was concentrated and Boc protection group was removed with the method described in general procedure for Boc deprotection. The crude product was purified with flash chromatography (CH_2Cl_2 – 5 % MeOH in CH_2Cl_2) to give the product.

General Procedure for the Syntheses of 4

Triphenylphosphine (524.6 mg, 2 mmol, 2.0 equiv.) was stirred in 2 mL dry CCl_4 at 55 ^oC for 3 h. Then 2 mL $CHCl_3$ was added, followed by **3a** (127.1 mg, 1.0 mmol, 1.0 equiv.). The resulting mixture was stirred at 55 ^oC under nitrogen for another 5 h. Solvent was removed under vacuum and the resulting dark red mixture was purified by flash chromatography (30 % ethyl acetate in hexanes) to give the product **4a** as a white crystal.



General Procedure for the Synthesis of 7

20 mL acetyl chloride was added dropwise into 20 mL MeOH at 0 °C over 30 min. After stirring for another 10 min, **5a** (1.84 g, 10 mmol) was added slowly as a 5 mL methanol solution. After the reaction mixture was stirred at room temperature for 12 h, 1 mL H_2O was added and stirred for 30 min. The solvent was removed under vacuum (note: the removal of solvents and HCl must be performed in a fume hood). Residual H_2O was azeotroped 3 times with toluene (3 × 30 mL) and the residue was placed under high vacuum for 3 h. (Note: When **5k'** or **5w** was the substrate, the reaction was performed at 0 °C for 12 h.)

Reductive amination reaction was performed with a modified literature procedure.[9] 100 mL 1,2-dichloroethane was added to the residue, followed by triethylamine (2.77 mL, 20 mmol, 2.0 equiv.) and then glacial acetic acid (0.58 mL, 10

mmol, 1.0 equiv.). **B** (benzyl 4-oxopiperidine-1-carboxylate) (2.33 g, 10 mmol, 1.0 equiv.) was added in one portion and the resulting mixture was stirred at room temperature for 20 min. The reaction was cooled to 0 ^{o}C and $NaBH(OAc)_3$ (5.30 g, 25 mmol, 2.5 equiv.) was added portionwise. Ice bath was removed and the reaction mixture was allowed to warm up to 30 ^{o}C and stirred for 3 h. CH_2Cl_2 (100 mL) was added and the mixture was washed with saturated $NaHCO_3$ solution 3 times (3 × 100 mL) (note: there will be a strong evolution of CO_2 in the first wash) and then with brine (100 mL). The organic layer was separated, dried over anhydrous $MgSO_4$, and concentrated in vacuum to give the product **6a** as colorless oil and it is used directly in the next step.

Bestmann ylide (recrystallized from toluene) (4.53 g, 15 mmol, 1.5 equiv.) was added to the solution of **6a** in 100 mL dry toluene at room temperature under N_2. The resulting mixture was stirred at 110 ^{o}C under N_2 for 48 h. After the reaction finished, toluene was removed under vacuum. 50 mL Et_2O was added and the resulting crude mixture was stirred at 0 ^{o}C for 30 min. The mixture was then filtered through a thin layer of celite to remove the solids. The filtered solution was concentrated under vacuum to give the product as oil. The oil was the mixture of the cyclized product **7a** and triphenylphosphine oxide and it was used without further purification in the next step for the syntheses of **8a** or **10a**.

Note: **B** (benzyl 4-oxopiperidine-1-carboxylate) was used for the syntheses of **6a**, **6t'**, **6s'**, **6v**, **6l**, **6w, 6f** and **C** (*tert*-butyl 4-oxopiperidine-1-carboxylate) was used for the syntheses of **6g**, **6k'**.

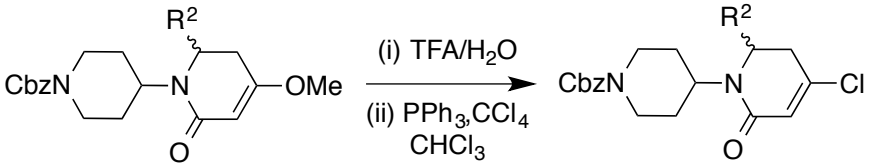
General Procedure for the Synthesis of 8

Syntheses of **8a**, **8t'**, **8l**, **8w**

General procedure for the deprotection of Cbz group was used to give the crude mixture. The crude product was purified by flash chromatography (3 % MeOH/ CH_2Cl_2 containing 1 % Et_3N - 8 % MeOH/CH_2Cl_2 containing 1 % Et_3N) to afford the product.

Syntheses Of **8g**, **8 k'**

General procedure for the deprotection of Boc group was used to give the crude mixture. The crude product was purified by flash chromatography (3 % MeOH/ CH_2Cl_2 containing 1 % Et_3N - 8 % MeOH/CH_2Cl_2 containing 1 % Et_3N) to afford the product.



7 **10**

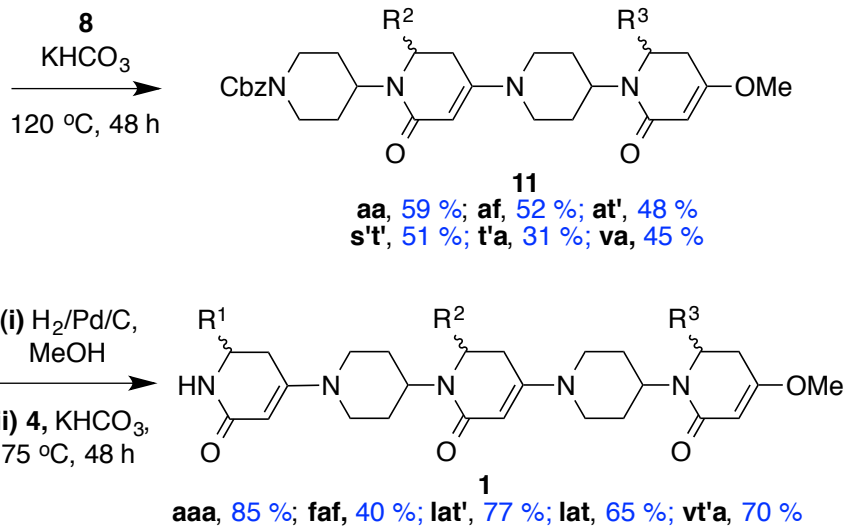
a, 41 %; **s'**, 45 %
t', 35 %; **v,** 31 %

General Procedure for the Syntheses of 10

7a (from 10 mmol **5a**) was dissolved in 50 mL 9:1 TFA:H_2O at 25 ^oC. After stirring for 18 h, toluene (50 mL) was added and the solution was concentrated. Residual TFA and water was azeotroped 3 times with toluene (50 mL × 5) and the residue was dissolved in 100 mL dichloromethane and the mixture was washed with 10 % $KHSO_4$ solution three times (3 × 50 mL) and then with brine (100 mL). The organic layer was

separated, dried over anhydrous $MgSO_4$, and concentrated in vacuum to give the product as colorless oil and it is used in the next step without further purification.

Triphenylphosphine (393 mg, 1.5 mmol, 1.5 equiv.) was stirred in 1.5 mL dry CCl_4 at 55 ^oC for 3 h. Then the piperidinedione product from the previous step was added as a 1.5 mL $CHCl_3$ solution. The resulting mixture was stirred at 55 ^oC under nitrogen for another 5 h. Solvent was removed under vacuum and the resulting dark red mixture was purified by flash chromatography (25 % ethyl acetate in hexanes) to give the product **10a** as colorless oil.



General Procedure for the Syntheses of 11

The general procedure for the coupling reaction was used at 120 ^oC for 48 h. The product was purified by flash chromatography (3 % $MeOH/CH_2Cl_2$ - 6 % $MeOH/CH_2Cl_2$) to afford the product.

General Procedure for the Syntheses of 1

General procedure for the deprotection of Cbz group was used to give the crude mixture of the deprotected product. The crude product was used directly in the next step without further purification.

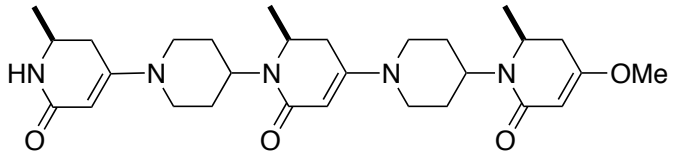
The Removal of –OBn Group for the Synthesis of 1lat

To a solution of deprotected two side-chain intermediate **at'** in ethanol (0.1 M) under nitrogen was added 1.0 equiv. of HCl in dioxane (4 M), followed by 10 wt % Pd/C (0.5 equiv. Pd). The reaction was placed under an atmosphere of hydrogen (1 atm, balloon) for 24 h. After the reaction finished, the flask was purged with N_2. The reaction mixture was filtered over a celite pad and concentrated to afford the crude product. The crude product was dissolved in dichloromethane and washed with saturated $NaHCO_3$ twice and then once with brine. The organic phase was dried over Na_2SO_4 and filtered. Dichloromethane was removed under vacuum to yield the desired product. The product was directly used in the next step.

The general procedure for the coupling reaction was used at 75 ^oC for 24 h. The product was purified by flash chromatography (5 % $MeOH/CH_2Cl_2$ - 8 % $MeOH/CH_2Cl_2$) to afford the product.

Compound **1aaa**

(*S*)-4-Methoxy-6-methyl-1-(1-((*S*)-2-methyl-1-(1-((*S*)-2-methyl-6-oxo-1,2,3,6-tetrahydropyridin-4-yl)piperidin-4-yl)-6-oxo-1,2,3,6-tetrahydropyridin-4-yl)piperidin-4-yl)-5,6-dihydropyridin-2(1*H*)-one



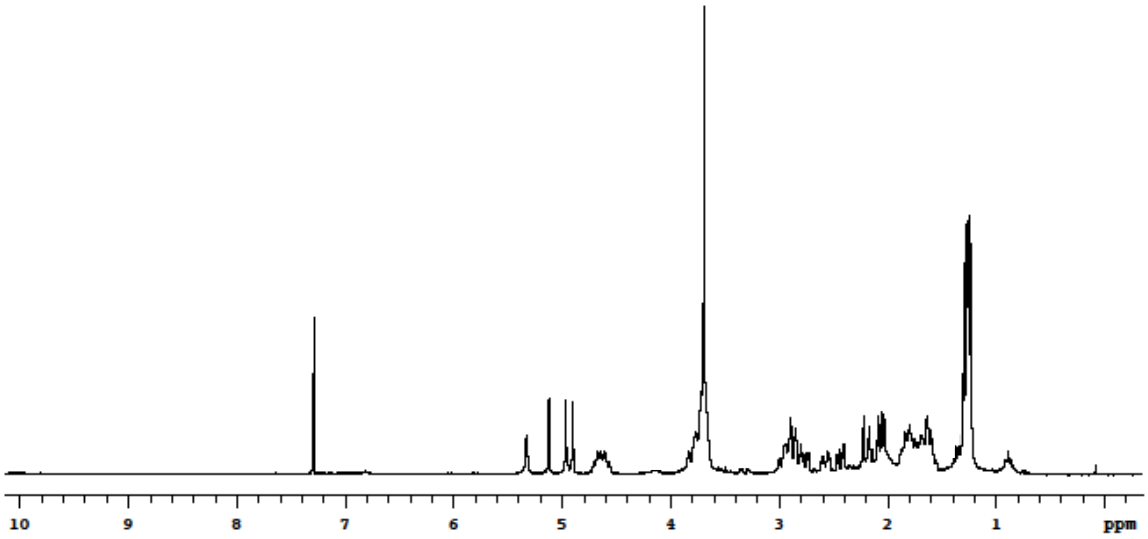
Light yellow oil, 85 %; $[\alpha]^{20}$ -30.7 (*c* 0.2, $CHCl_3$)

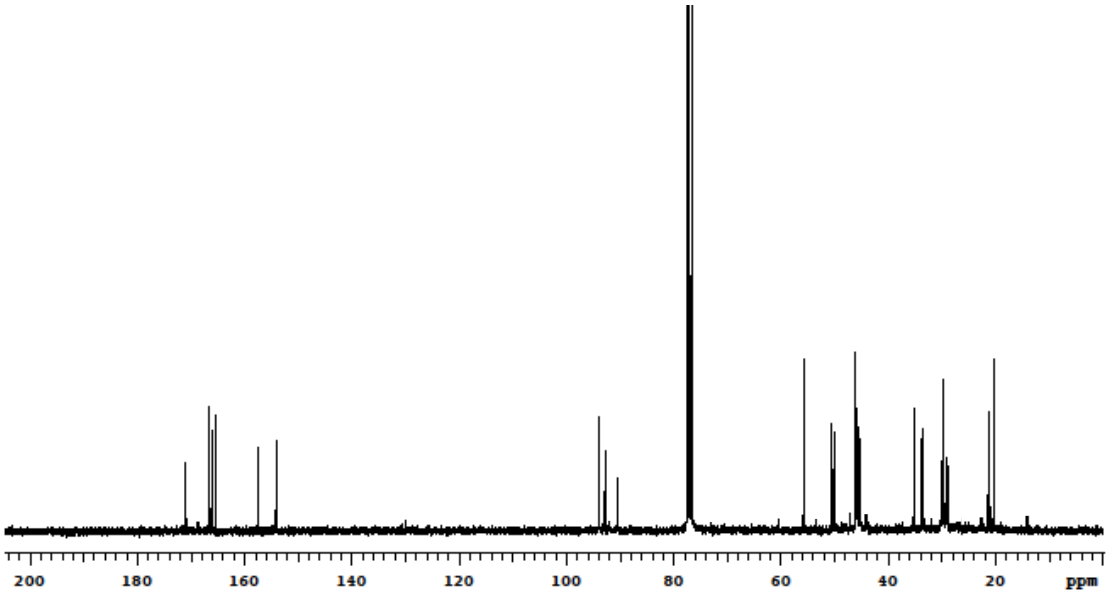
^{1}H-NMR (300 MHz, $CDCl_3$) δ 5.33 (s, 1H), 5.12 (d, *J* = 1.8 Hz, 1H), 4.97 (s, 1H), 4.90 (s, 1H), 4.76-4.57 (m, 2H), 3.84-3.61 (m, 10H), 3.01-2.69 (m, 6H), 2.58-2.50 (m, 1H), 2.49-2.39 (m, 1H), 2.12-1.98 (m, 2H), 1.90-1.57 (m, 8H), 1.34-1.27 (m, 9H);

^{13}C-NMR (75 MHz, $CDCl_3$) δ 170.9, 166.5, 165.9, 165.4, 157.5, 153.8, 94.0, 92.8, 90.5, 55.6, 50.5, 49.9, 46.3, 46.2, 46.1, 45.9, 45.4, 45.2, 45.1, 35.3, 34.0, 33.4, 30.2, 29.9, 29.1, 28.9, 21.2, 20.4, 20.3;

IR (film, cm^{-1}) 3406, 2927, 2361, 1603, 1433, 1377, 1321, 1276, 1207, 1175, 1005, 808, 731;

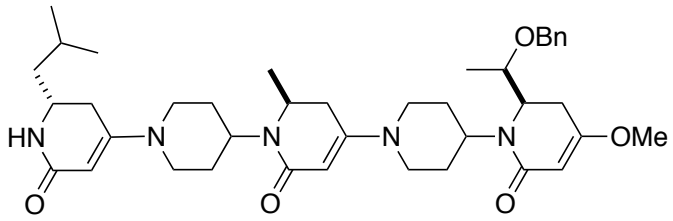
HRMS (ESI) m/z calcd for $C_{29}H_{43}N_5O_4$ $(M+H)^+$ 526.3393; found 526.3411 (3.4 ppm).





Compound **1lat'**

(*R*)-6-((*R*)-1-(Benzyloxy)ethyl)-1-(1-((*S*)-1-(1-((*R*)-2-isobutyl-6-oxo-1,2,3,6-tetrahydropyridin-4-yl)piperidin-4-yl)-2-methyl-6-oxo-1,2,3,6-tetrahydropyridin-4-yl)piperidin-4-yl)-4-methoxy-5,6-dihydropyridin-2(1*H*)-one



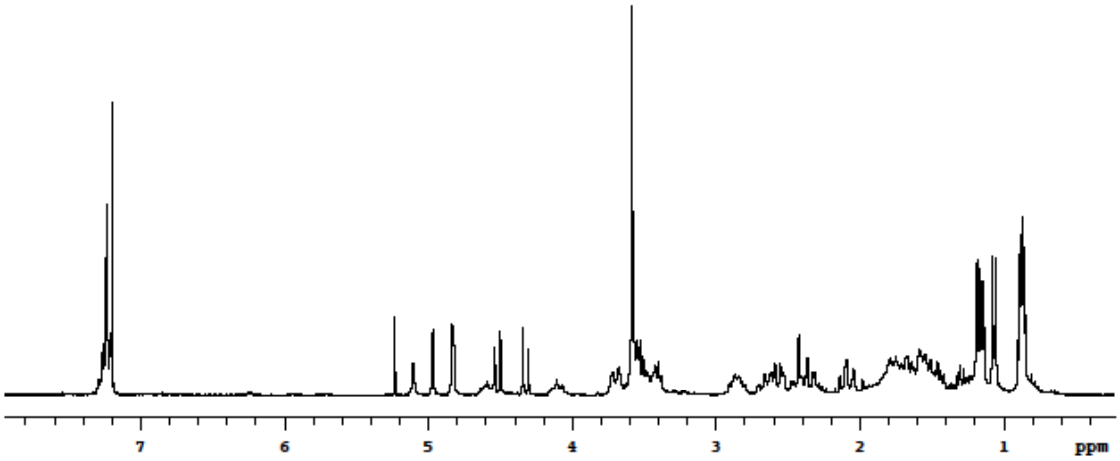
Colorless oil, 77 %; $[\alpha]^{20}$ -21.5 (*c* 0.1, $CHCl_3$)

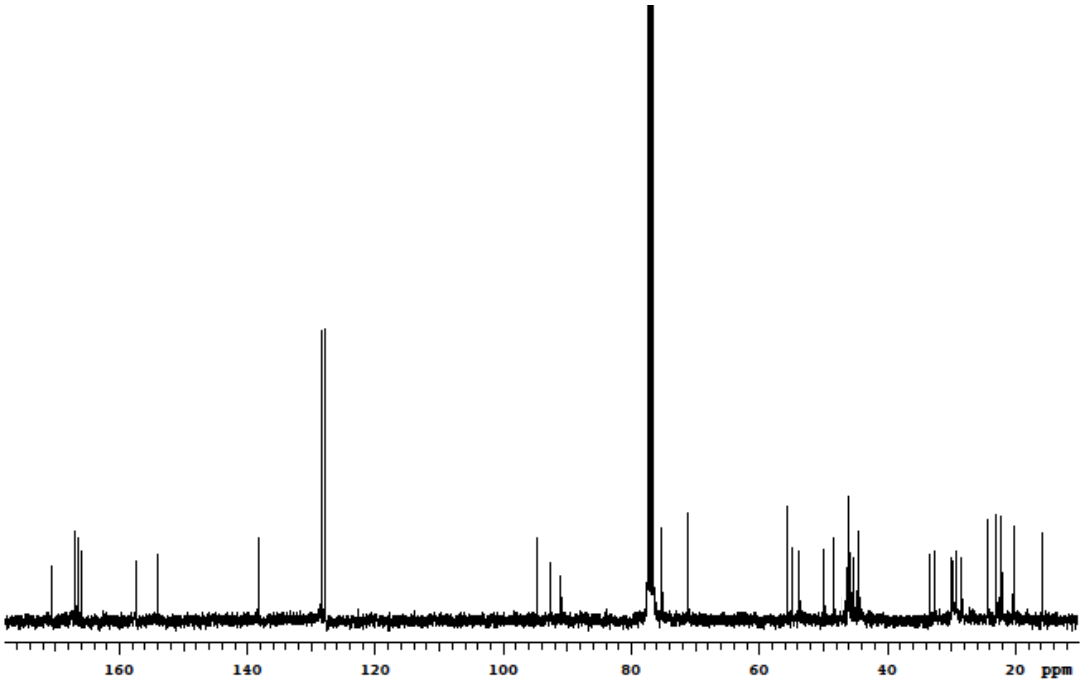
^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.29-7.20 (m, 5H), 5.10 (s, 1H), 4.97 (s, 1H), 4.83 (s, 2H), 4.65-4.47 (m, 2H), 4.32 (d, *J* = 11.7 Hz, 1H), 4.18-4.03 (m, 1H), 3.76-3.38 (m, 11H), 2.94-2.75 (m, 2H), 2.70-2.24 (m, 6H), 2.18-2.01 (m, 2H), 1.92-1.41 (m, 10H), 1.37-1.21 (m, 1H), 1.29 (d, *J* = 6.3 Hz, 3H), 1.07 (d, *J* = 6.3 Hz, 3H), 0.90-0.84 (m, 6H);

^{13}C-NMR (75 MHz, $CDCl_3$) δ 170.7, 166.9, 166.5, 165.9, 157.4, 154.1, 138.1, 128.4, 127.8, 127.7, 94.8, 92.6, 91.0, 75.2, 71.1, 55.7, 55.0, 53.8, 49.9, 48.3, 46.2, 46.0, 45.8, 45.2, 44.4, 33.4, 32.7, 30.0, 29.8, 29.7, 29.4, 29.2, 28.4, 24.3, 23.0, 22.1, 20.3, 15.9;

IR (film, cm^{-1}) 2925, 1614, 1428, 1275, 1224, 1076, 807;

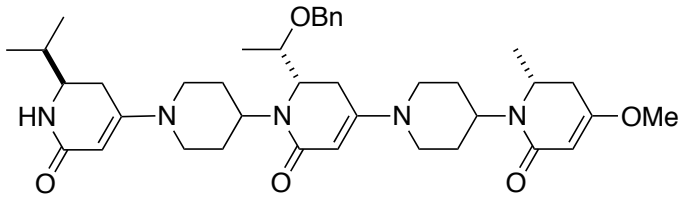
HRMS (ESI) m/z calcd for $C_{40}H_{57}N_5O_5$ $(M+H)^+$ 688.4438; found 688.4463 (3.6 ppm).





Compound **1vt'a**

(*S*)-6-((*S*)-1-(Benzyloxy)ethyl)-1-(1-((*R*)-2-isopropyl-6-oxo-1,2,3,6-tetrahydropyridin-4-yl)piperidin-4-yl)-4-(4-((*R*)-4-methoxy-6-methyl-2-oxo-5,6-dihydropyridin-1(2*H*)-yl)piperidin-1-yl)-5,6-dihydropyridin-2(1*H*)-one



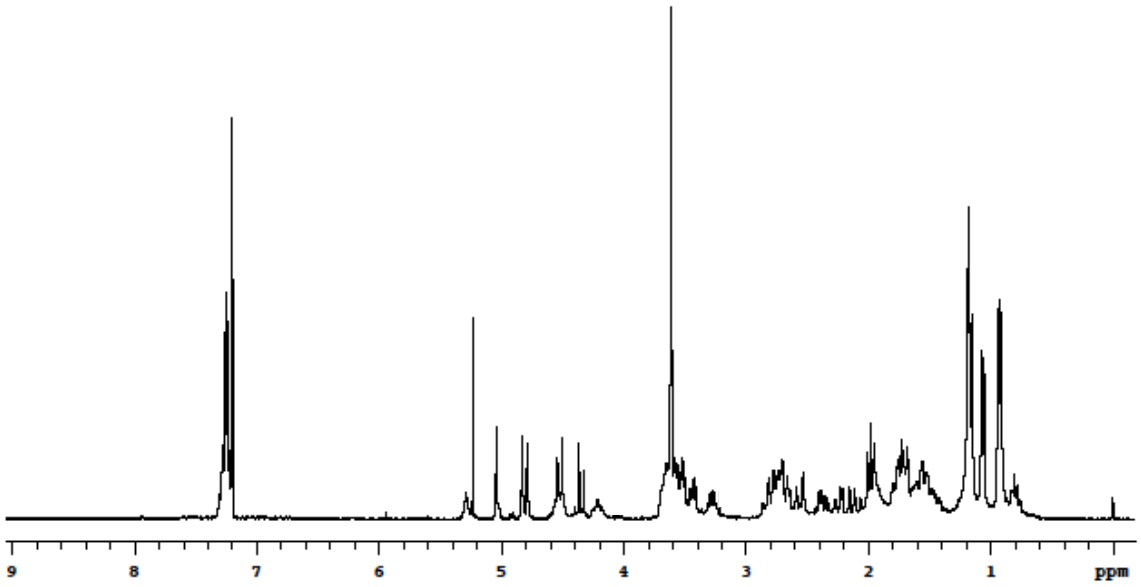
Colorless oil, 70 %; $[\alpha]^{20}$ -32.4 (*c* 0.2, $CHCl_3$)

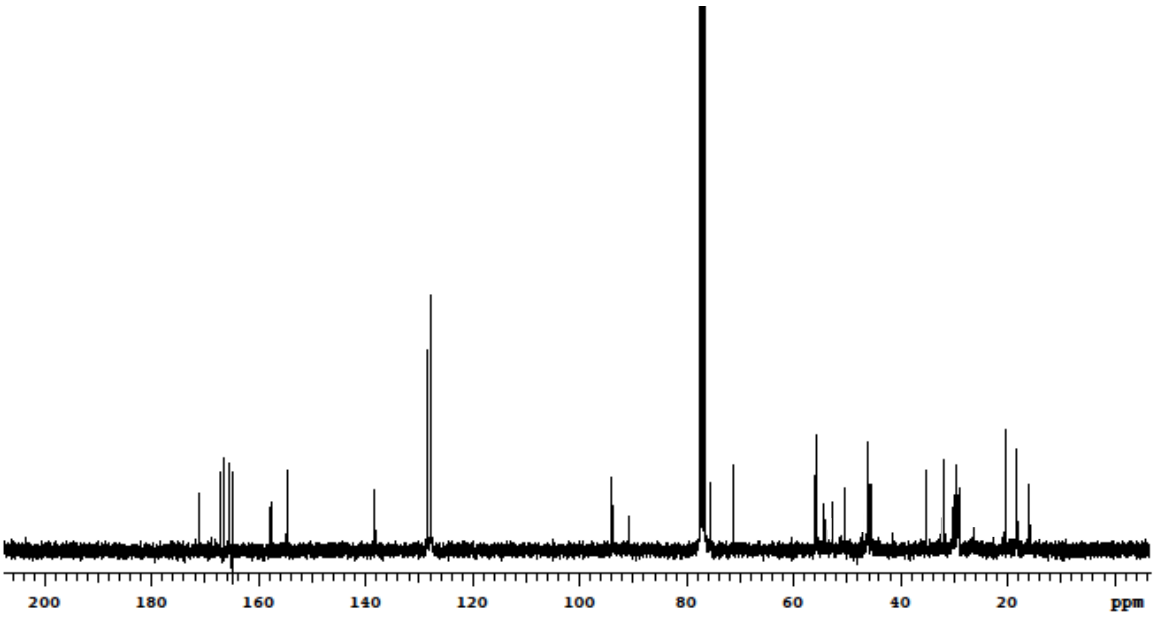
^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.30-7.20 (m, 5H), 5.29 (br, 1H), 5.04 (s, 1H), 4.83 (s, 1H), 4.79 (s, 1H), 4.59-4.44 (m, 2H), 4.34 (d, *J* = 12.0 Hz, 1H), 4.26-4.10 (m, 1H), 3.73-3.22 (m, 11H), 2.89-2.51 (m, 5H), 2.42-1.84 (m, 5H), 1.82-1.38 (m, 9H), 1.17 (d, *J* = 6.6 Hz, 3H), 1.07 (d, *J* = 6.3 Hz, 3H), 0.95-0.90 (m, 6H);

^{13}C-NMR (75 MHz, $CDCl_3$) δ 171.0, 167.2, 166.5, 165.4, 157.7, 154.4, 138.2, 128.5, 127.8, 127.7, 94.1, 93.7, 90.8, 75.6, 71.2, 56.1, 55.6, 54.5, 52.8, 50.5, 46.2, 46.1, 46.0, 45.9, 45.4, 35.3, 31.8, 30.3, 30.0, 29.7, 29.4, 29.2, 29.1, 20.5, 18.4, 18.3, 15.8;

IR (film, cm^{-1}) 3406 (br), 2927, 2360, 1614, 1433, 1385, 1274, 1243, 1207, 1174, 1078, 1008, 818, 733;

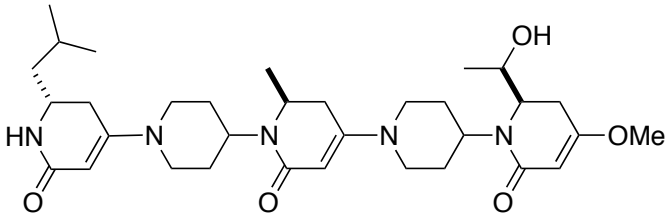
HRMS (ESI) m/z calcd for $C_{39}H_{55}N_5O_5$ $(M+H)^+$ 674.4281; found 674.4301 (3.0 ppm).





Compound **1lat**

(*R*)-6-((*R*)-1-Hydroxyethyl)-1-(1-((*S*)-1-(1-((*R*)-2-isobutyl-6-oxo-1,2,3,6-tetrahydropyridin-4-yl)piperidin-4-yl)-2-methyl-6-oxo-1,2,3,6-tetrahydropyridin-4-yl)piperidin-4-yl)-4-methoxy-5,6-dihydropyridin-2(1*H*)-one



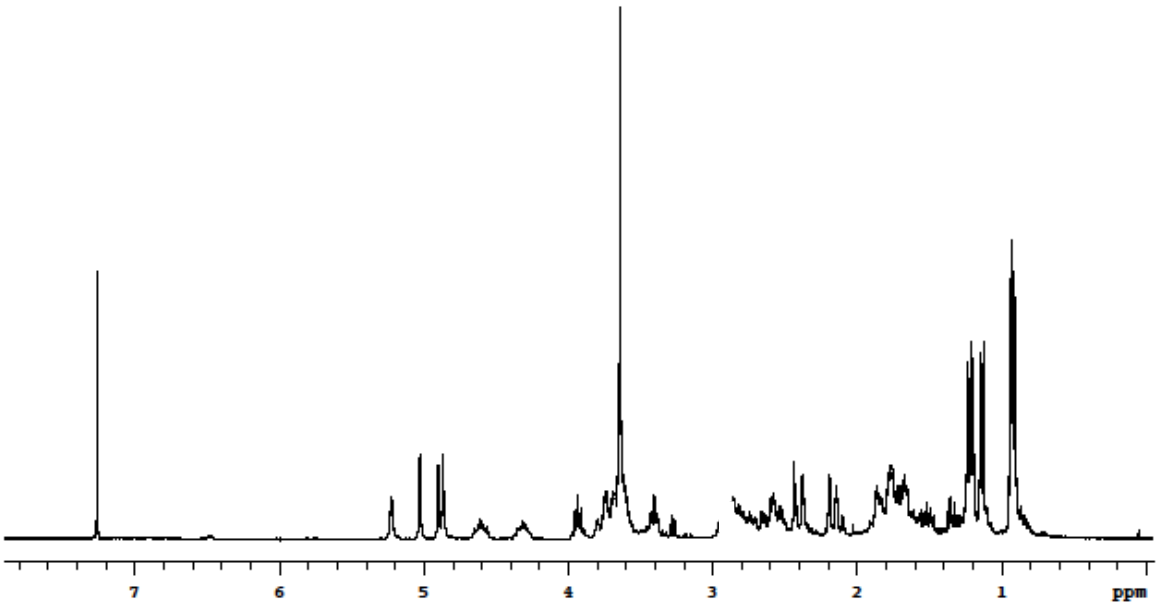
White solid, 65 %; mp = 122.4 – 124.2 °C; $[\alpha]^{20}$ +15.7 (*c* 0.2, $CHCl_3$)

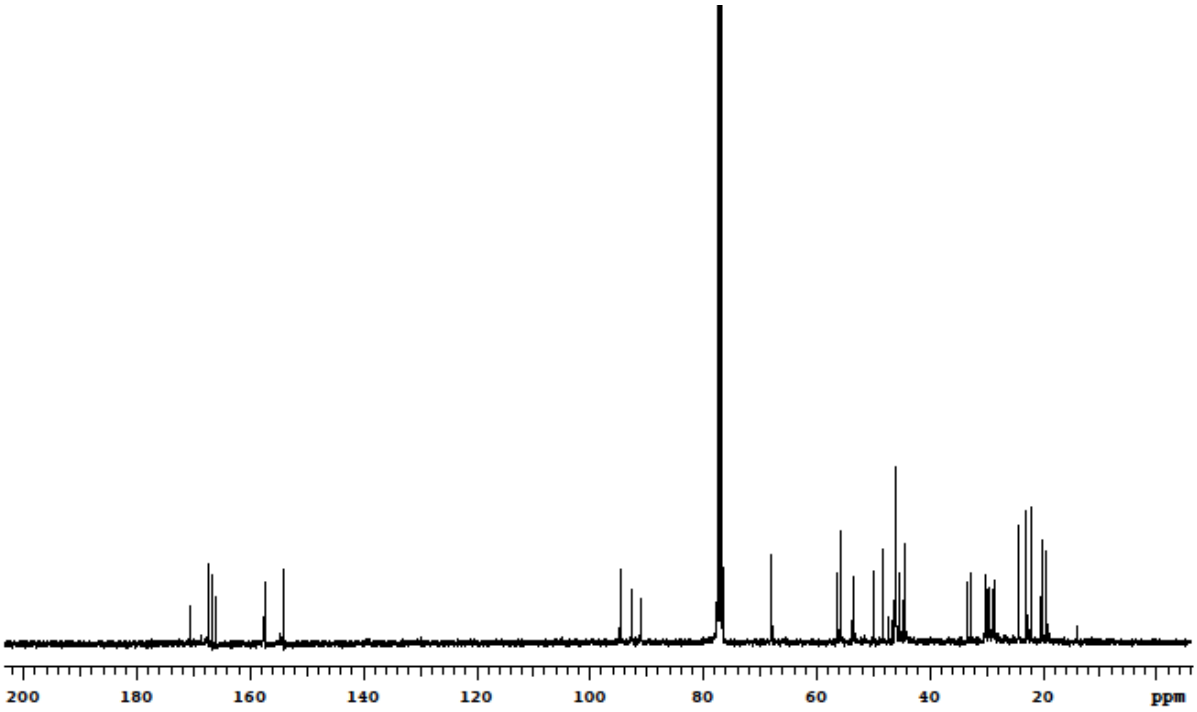
^{1}H-NMR (300 MHz, $CDCl_3$) δ 5.03 (d, *J* = 1.5 Hz, 1H), 4.90 (s, 1H), 4.87 (s, 1H), 4.66-4.54 (m, 1H), 4.39-4.23 (m, 1H), 3.99-3.85 (m, 1H), 3.82-3.54 (m, 9H), 3.47-3.38 (m, 1H), 2.97-2.43 (m, 6H), 2.42-2.38 (m, 2H), 2.20-2.14 (m, 2H), 1.90-1.44 (m, 11H), 1.38-1.25 (m, 1H), 1.21 (d, *J* = 6.6 Hz, 3H), 1.13 (d, *J* = 6.3 Hz, 3H), 0.96-0.91 (m, 6H);

^{13}C-NMR (75 MHz, $CDCl_3$) δ 170.9, 167.2, 166.7, 166.0, 157.5, 154.1, 94.6, 92.5, 91.0, 67.9, 56.3, 55.7, 53.6, 50.0, 48.3, 46.1, 46.1, 45.9, 45.3, 44.4, 33.4, 32.7, 30.2, 30.1, 29.7, 29.5, 28.9, 28.5, 24.3, 23.0, 22.2, 20.3, 19.4;

IR (film, cm^{-1}) 3286, 2927, 1601, 1432, 1383, 1277, 1225, 1078, 1007, 804;

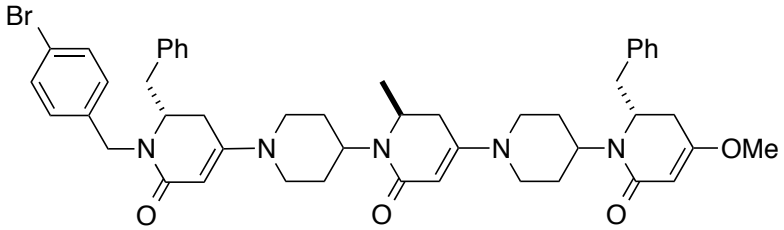
HRMS (ESI) m/z calcd for $C_{33}H_{51}N_5O_5$ $(M+H)^+$ 598.3968; found 598.3941 (4.3 ppm).


7
6
5
4
3
2
1
ppm


200
180
160
140
120
100
80
60
40
20
ppm

Compound **1faf**

(*R*)-6-benzyl-1-(1-((*S*)-1-(1-((*R*)-2-benzyl-1-(4-bromobenzyl)-6-oxo-1,2,3,6-tetrahydropyridin-4-yl)piperidin-4-yl)-2-methyl-6-oxo-1,2,3,6-tetrahydropyridin-4-yl)piperidin-4-yl)-4-methoxy-5,6-dihydropyridin-2(1*H*)-one

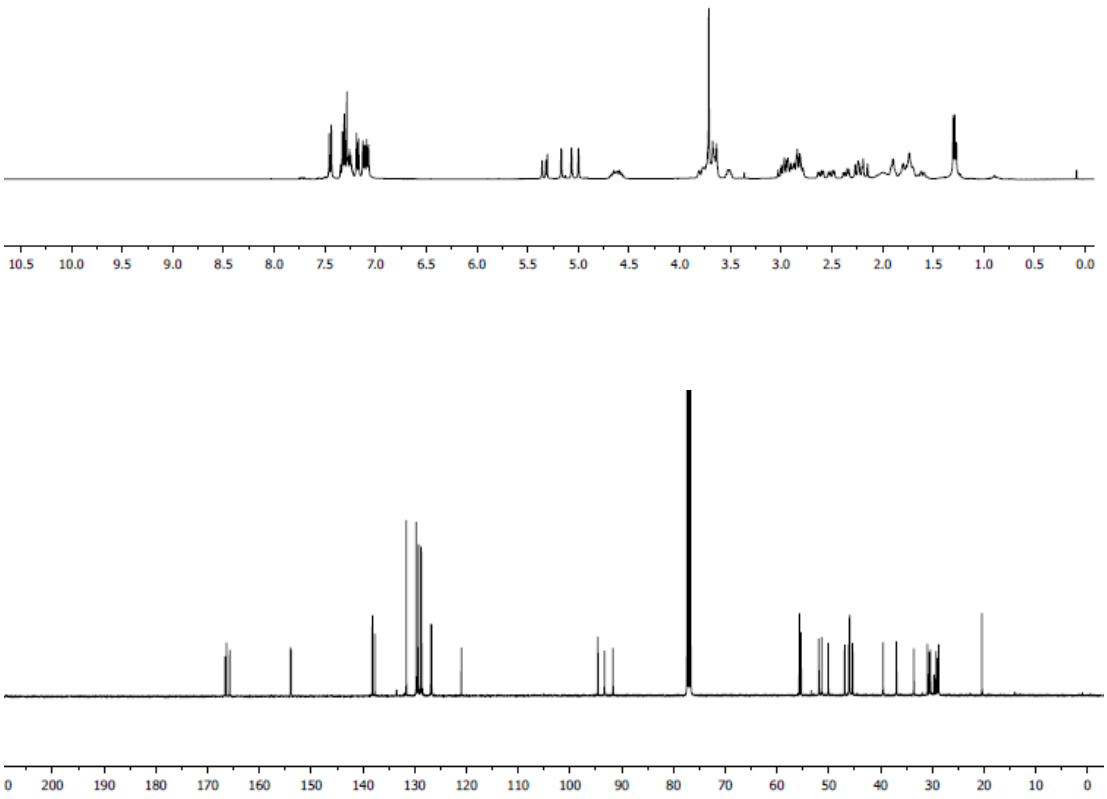


White solid, 40 %; mp = 135.1 – 136.4 °C; $[\alpha]^{20}$ -23.3 (*c* 0.2, $CHCl_3$);

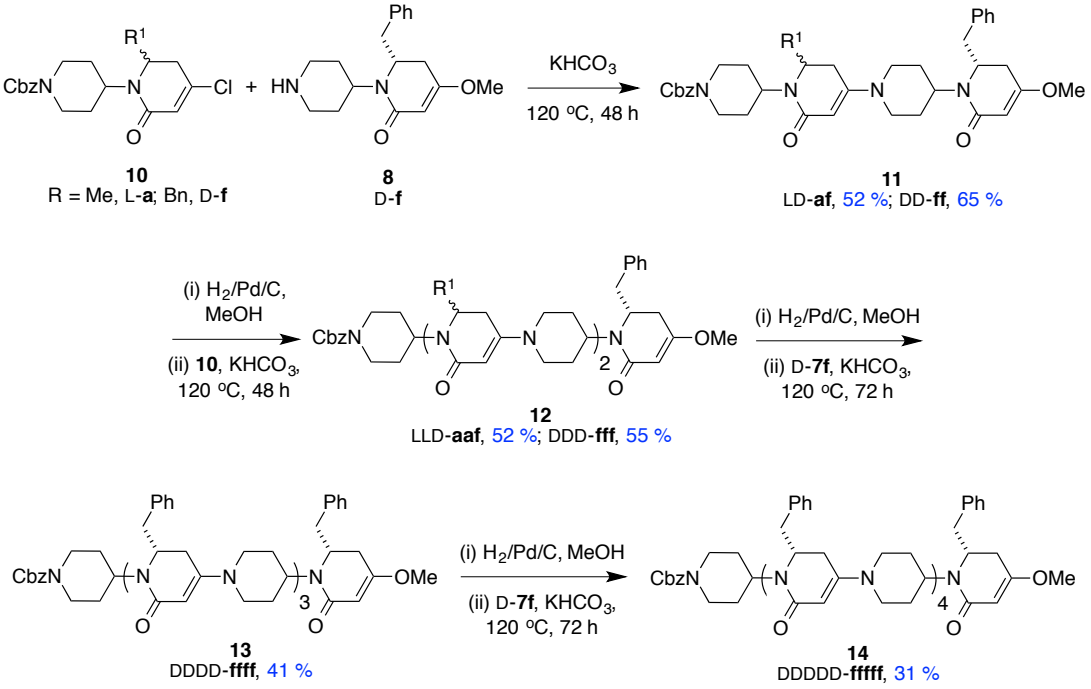
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.45 (d, *J* = 8.3 Hz, 2H), 7.39-7.20 (m, 6H), 7.18 (d, *J* = 8.3 Hz, 1H), 7.16-7.04 (m, 4H), 5.34 (d, *J* = 15.3 Hz, 1H), 5.17 (d, *J* = 1.4 Hz, 1H), 5.07 (s, 1H), 5.00 (s, 1H), 4.70 – 4.53 (m, 2H), 3.83 – 3.60 (m, 10H), 3.58-3.48 (m, 1H), 3.05 – 2.76 (m, 8H), 2.61 (dd, *J* = 15.5, 5.9 Hz, 1H), 2.50 (dd, *J* = 16.9, 5.7 Hz, 1H), 2.36 (dd, *J* = 15.8, 6.1 Hz, 1H), 2.29 – 2.13 (m, 3H), 2.07 – 1.52 (m, 8H), 1.30 (d, *J* = 6.5 Hz, 3H);

^{13}C NMR (101 MHz, $CDCl_3$) δ 166.64, 166.39, 165.77, 165.73, 153.97, 153.86, 138.17, 138.16, 137.69, 131.62, 129.72, 129.31, 129.24, 128.77, 128.68, 126.89, 126.72, 120.98, 94.60, 93.32, 91.70, 55.68, 55.34, 51.91, 51.35, 50.10, 46.92, 46.12, 46.01, 45.97, 45.39, 39.50, 36.92, 33.57, 30.92, 30.57, 30.35, 29.66, 29.28, 28.95, 28.77, 20.41;

HRMS (ESI) m/z calcd for $C_{48}H_{56}BrN_5O_4$ $(M+H)^+$ 846.3594; found 846.3564 (3.5 ppm).



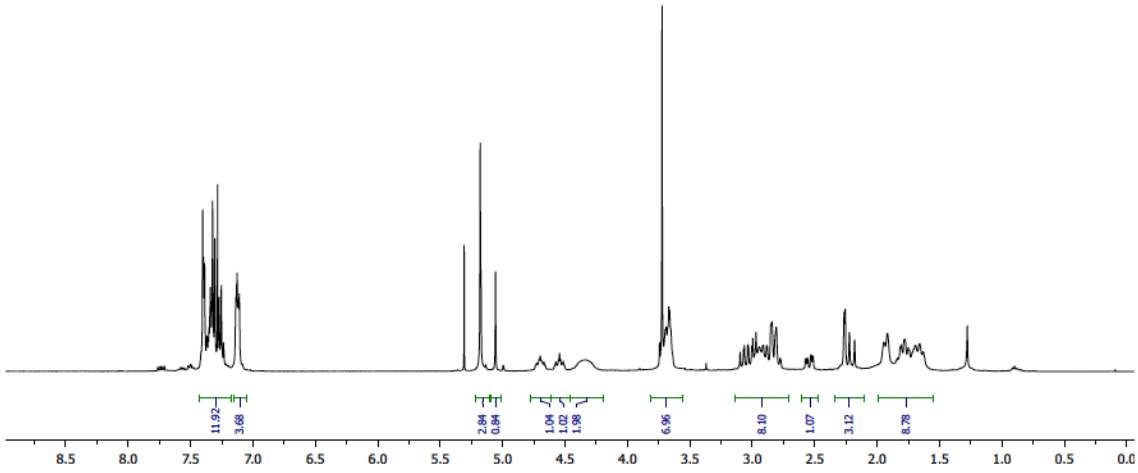
C. Synthesis of Oligomeric Mimic 14

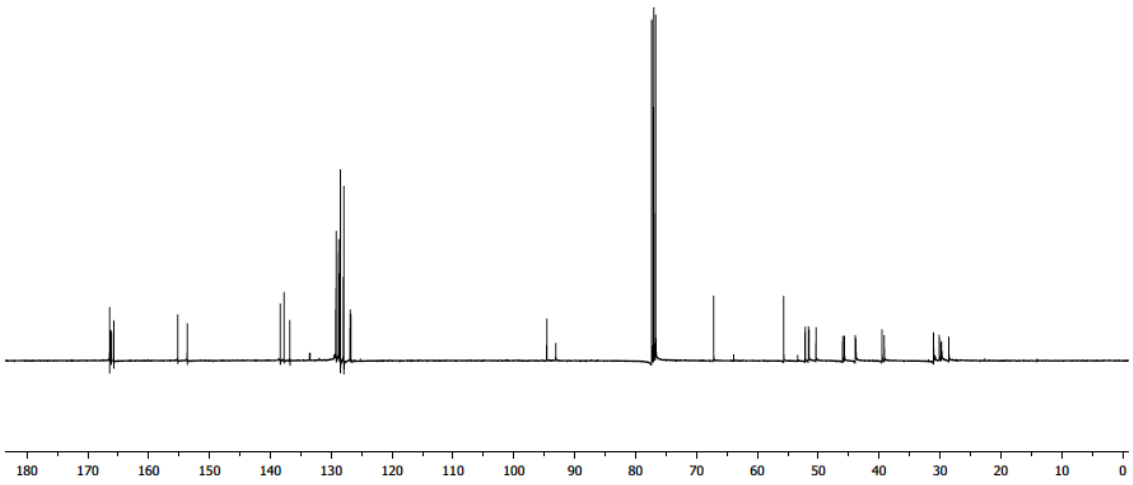


Anhydrous $KHCO_3$ powder (4.00 g, 40.0 mmol, 40.0 equiv.) was added to a 0.5 M solution of D-**8f** (300.5 mg, 1.0 mmol, 1.0 equiv.) in CH_2Cl_2. The vinyl chloride D-**10f** (439.0 mg, 1.0 mmol, 1.0 equiv.) was then added as a 0.5 M solution in CH_2Cl_2 to the reaction mixture. While stirring at room temperature, the solvent was removed with nitrogen. The resulting solid was heated up to the 120 ºC under N_2. After 24 h, 2 mL anhydrous 1,4-dioxane was added, stirred for 10 mins and then removed with nitrogen. The mixture was kept at 120 ºC under N_2 for another 24 h. After the reaction finished, the mixture was cooled down to room temperature. 20 mL CH_2Cl_2 was added and the inorganic solids were removed by filtration. The solution was concentrated to give the

crude product. The product was purified by flash chromatography (SiO_2, 3 % MeOH/CH_2Cl_2 - 5 % MeOH/CH_2Cl_2) to afford the product DD-**11ff**.

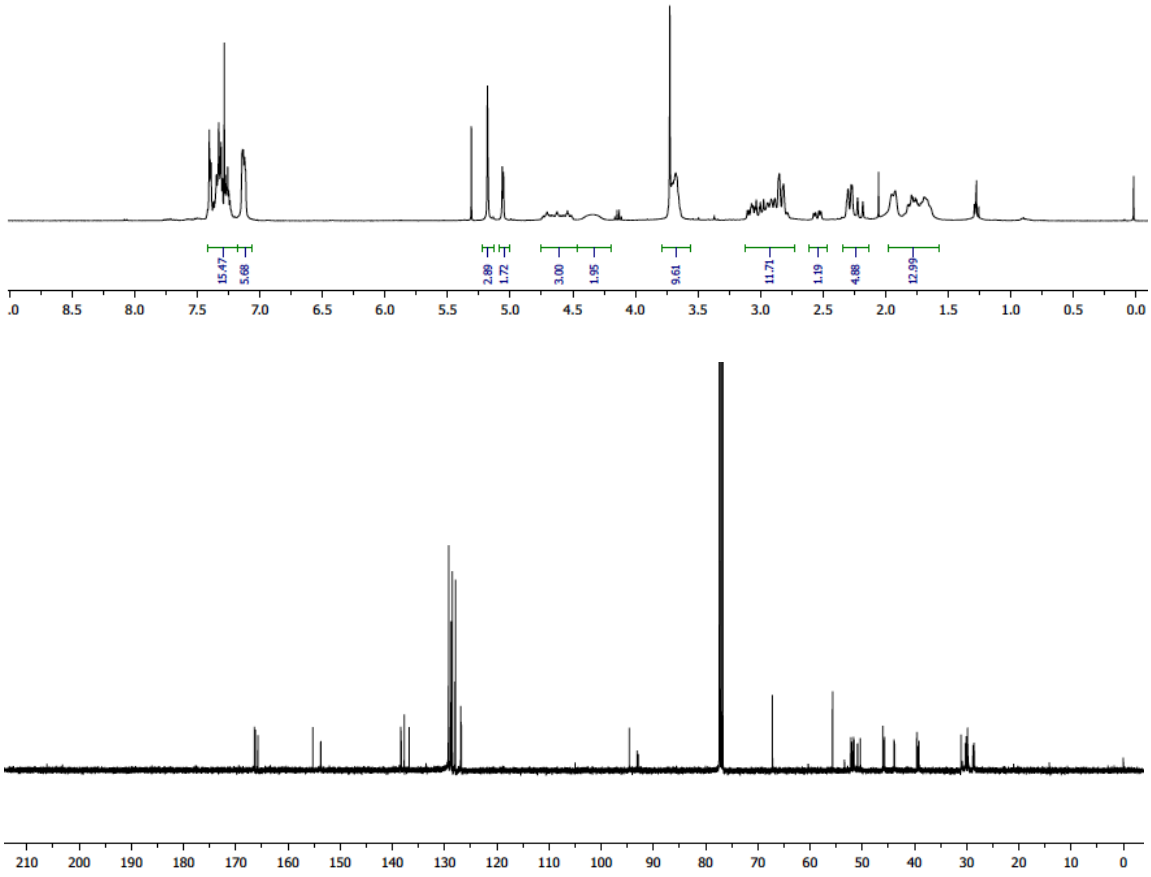
Benzyl 4-((*R*)-2-benzyl-4-(4-((*R*)-2-benzyl-4-methoxy-6-oxo-3,6-dihydropyridin-1(2*H*)-yl)piperidin-1-yl)-6-oxo-3,6-dihydropyridin-1(2*H*)-yl)piperidine-1-carboxylate (DD-**11ff**): white foamy solid, 454.3 mg from coupling of D-**11f** and D-**8f** with general procedure A, 65 % yield; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.40-7.20 (m, 12H), 7.18 – 7.04 (m, 3H), 5.16 (s, 3H), 5.06 (s, 1H), 4.77 – 4.61 (m, 1H), 4.59-4.50 (m, 1H), 4.33 (br, 2H), 3.75 – 3.60 (m, 7H), 3.14 – 2.71 (m, 8H), 2.54 (dd, *J* = 16.9, 5.6 Hz, 1H), 2.34 – 2.09 (m, 3H), 1.99 – 1.55 (m, 8H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 166.42, 166.17, 165.75, 155.21, 153.67, 138.38, 137.75, 136.80, 129.26, 129.18, 128.79, 128.70, 128.51, 128.03, 127.93, 126.91, 126.80, 94.60, 93.12, 67.19, 55.69, 52.16, 51.62, 51.53, 50.38, 45.99, 45.71, 43.91, 43.82, 39.54, 39.16, 31.06, 30.83, 30.11, 29.79, 29.67, 28.57; HRMS (ESI+): m/z calcd for $C_{43}H_{50}N_4O_5$ $[M+H]^+$ 703.3859; found 703.384.





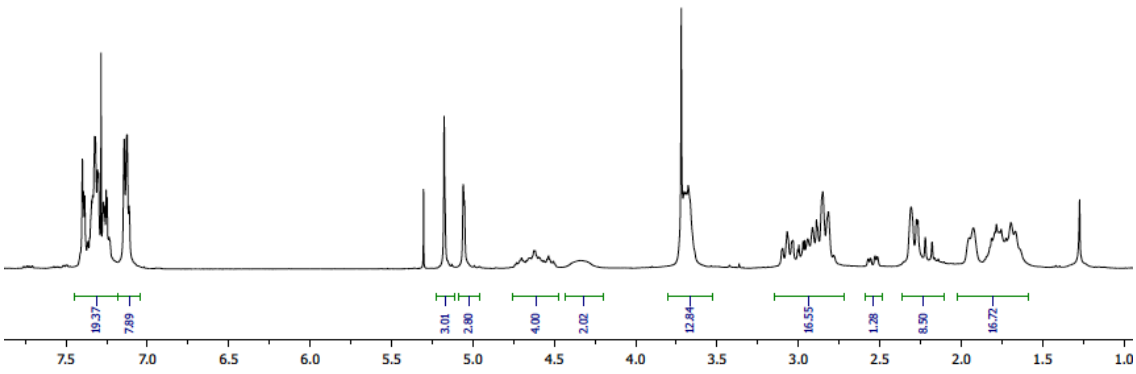
Benzyl 4-((*R*)-2-benzyl-4-(4-((*R*)-2-benzyl-4-(4-((*R*)-2-benzyl-4-methoxy-6-oxo-3,6-dihydropyridin-1(2*H*)-yl)piperidin-1-yl)-6-oxo-3,6-dihydropyridin-1(2*H*)-yl)piperidin-1-yl)-6-oxo-3,6-dihydropyridin-1(2*H*)-yl)piperidine-1-carboxylate (DDD-**12fff**). DD-**11ff** (351.3 mg, 0.5 mmol) was deprotected with general procedure B and the product was coupled with D-**10f** (220.5 mg, 0.5 mmol) using general procedure for coupling reaction to give DDD-**12fff** as a light yellow oil. 268.8 mg, 55 % yield; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.43 – 7.05 (m, 20H), 5.18 (s, 3H), 5.05 (s, 1H), 5.06 (s, 1H), 4.76 – 4.48 (m, 3H), 4.35 (br, 2H), 3.74 – 3.63 (m, 10H), 3.12 – 2.76 (m, 12H), 2.55 (dd, *J* = 16.8, 5.6 Hz, 1H), 2.33-2.13 (m, 5H), 2.00 – 1.61 (m, 12H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 166.42, 166.25, 165.76, 155.22, 153.72, 153.67, 138.41, 138.32, 137.74, 136.80, 129.28, 129.19, 128.80, 128.75, 128.70, 128.51, 128.03, 127.93, 126.92, 126.87, 126.79, 94.60, 93.11, 92.90, 67.20, 55.69, 52.20, 51.97, 51.66, 51.56, 50.91, 50.38, 46.02, 45.77, 45.74, 43.92, 43.83, 39.54, 39.24, 39.16, 31.06, 30.84, 30.23, 30.09, 29.81,

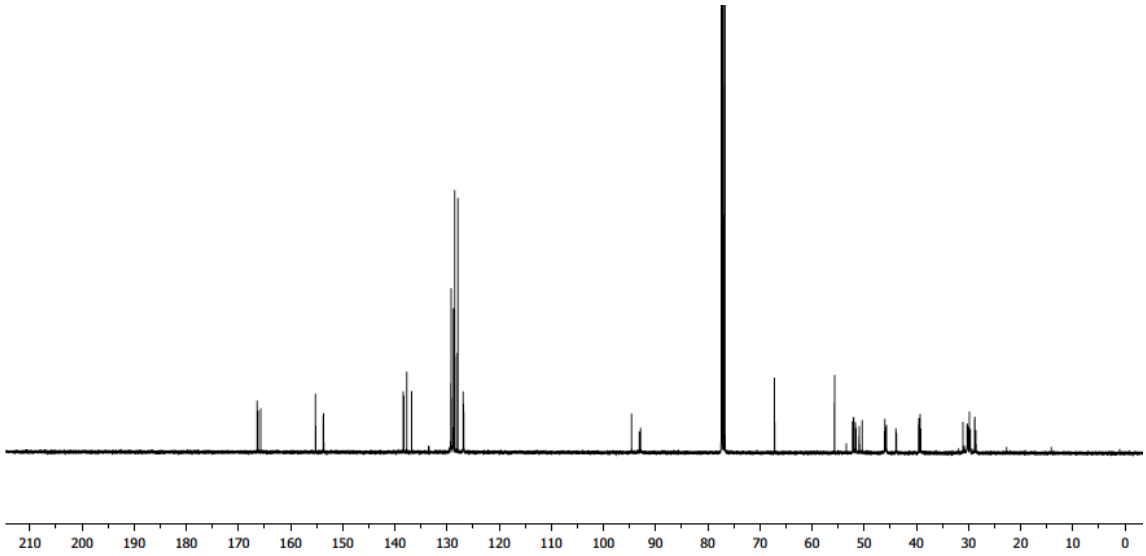
29.67, 28.78, 28.56; HRMS (ESI+): m/z calcd for $C_{60}H_{70}N_6O_6$ $[M+H]^+$ 971.5435; found 971.5437.



Benzyl 4-((*R*)-2-benzyl-4-(4-((*R*)-2-benzyl-4-(4-((*R*)-2-benzyl-4-(4-((*R*)-2-benzyl-4-methoxy-6-oxo-3,6-dihydropyridin-1(2*H*)-yl)piperidin-1-yl)-6-oxo-3,6-dihydropyridin-1(2*H*)-yl)piperidin-1-yl)-6-oxo-3,6-dihydropyridin-1(2*H*)-yl)piperidin-1-yl)-6-oxo-3,6-dihydropyridin-1(2*H*)-yl)piperidine-1-carboxylate (DDDD-**13ffff**). DDD-**12fff** (194.9 mg, 0.2 mmol) was deprotected with general procedure for Cbz

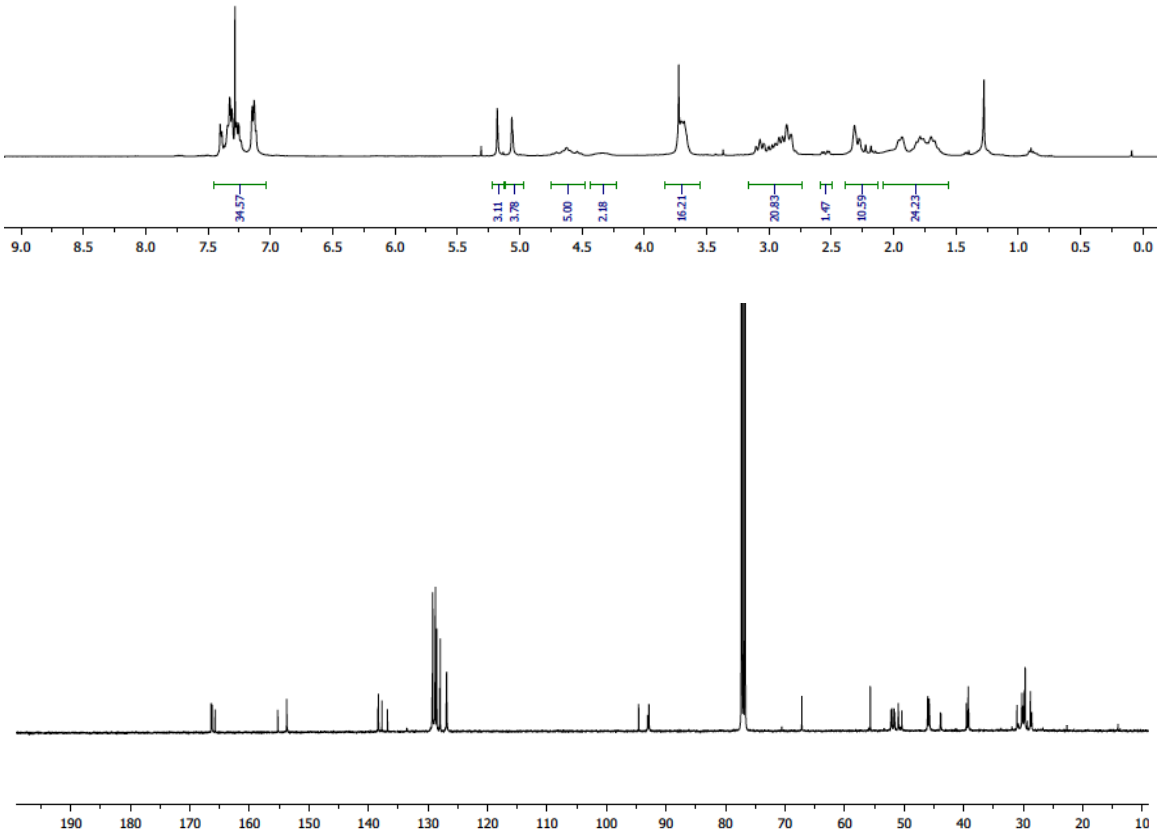
deprotection and the product was coupled with D-**10f** (175.6 mg, 0.4 mmol) using general procedure for coupling reaction to give DDDD-**13ffff** as a light yellow oil. 99.6 mg, 41 % yield; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.44 – 7.08 (m, 25H), 5.19 (s, 3H), 5.10 – 5.02 (m, 3H), 4.79 – 4.47 (m, 4H), 4.35 (br, 2H), 3.80 – 3.61 (m, 13H), 3.14 – 2.76 (m, 16H), 2.55 (dd, *J* = 16.8, 5.8 Hz, 1H), 2.39 – 2.16 (m, 7H), 2.03 – 1.57 (m, 16H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 166.41, 166.31, 166.24, 165.76, 155.23, 153.71, 153.66, 138.42, 138.36, 138.32, 137.75, 136.80, 129.29, 129.20, 128.80, 128.75, 128.70, 128.51, 128.03, 127.94, 126.92, 126.88, 126.79, 94.62, 93.15, 92.93, 92.92, 67.20, 55.69, 52.21, 52.02, 51.66, 51.57, 50.95, 50.91, 50.36, 46.06, 46.03, 45.80, 45.75, 43.92, 43.83, 39.54, 39.25, 39.17, 31.06, 30.84, 30.25, 30.10, 29.94, 29.82, 29.68, 28.79, 28.57; HRMS (MALDI+): m/z calcd for $C_{77}H_{90}N_8O_7$ $[M+Na]^+$ 1261.6824; found 1261.6881.





Benzyl 4-((*R*)-2-benzyl-4-(4-((*R*)-2-benzyl-4-(4-((*R*)-2-benzyl-4-(4-((*R*)-2-benzyl-4-(4-((*R*)-2-benzyl-4-methoxy-6-oxo-3,6-dihydropyridin-1(2*H*)-yl)piperidin-1-yl)-6-oxo-3,6-dihydropyridin-1(2*H*)-yl)piperidin-1-yl)-6-oxo-3,6-dihydropyridin-1(2*H*)-yl)piperidin-1-yl)-6-oxo-3,6-dihydropyridin-1(2*H*)-yl)piperidin-1-yl)-6-oxo-3,6-dihydropyridin-1(2*H*)-yl)piperidine-1-carboxylate (DDDDD-**14fffff**): DDDD-**13ffff** (75.1 mg, 0.06 mmol) was deprotected with general procedure B and the product was coupled with D-**10f** (88.5 mg, 0.2 mmol) using general procedure A to give DDDDD-**14fffff** as a light yellow oil. 27.6 mg, 31 % yield; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.49 – 7.07 (m, 30H), 5.21-5.17 (m, 3H), 5.10 – 5.02 (m, 4H), 4.77 – 4.49 (m, 5H), 4.35 (br, 2H), 3.77 – 3.62 (m, 16H), 3.15 – 2.79 (m, 20 H), 2.56 (dd, *J* = 16.9, 5.6 Hz, 1H), 2.37 – 2.15 (m, 9H), 2.09 – 1.61 (m, 20H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 166.42, 166.33, 166.26, 165.77, 155.22, 153.73, 153.67, 138.40, 138.34, 138.31, 137.73, 136.79, 129.28, 129.19, 128.80, 128.75, 128.69, 128.50, 128.03, 127.93, 126.91, 126.87, 126.79, 94.60, 93.10,

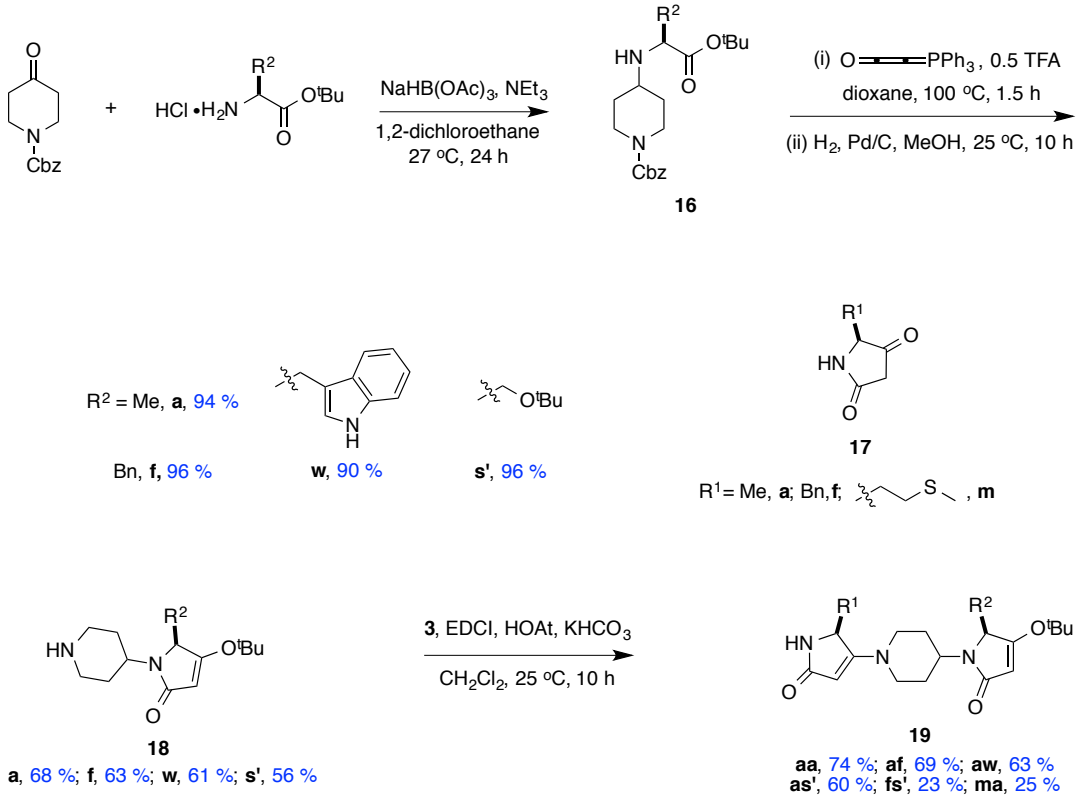
92.87, 67.19, 55.69, 52.22, 52.02, 51.69, 51.56, 50.98, 50.93, 50.38, 46.03, 45.81, 45.74, 43.91, 43.83, 39.53, 39.24, 39.16, 31.06, 30.23, 30.09, 29.81, 29.67, 28.79, 28.56; HRMS (MALDI+): m/z calcd for $C_{94}H_{110}N_{10}O_8$ $[M+Na]^+$ 1529.8398; found 1529.8406.



APPENDIX C

EXPERIMENTAL PROCEDURES FOR CHAPTER III

A. Synthesis of Oligo-piperidines-pyrrolidinones



General Procedure for the Syntheses of 18

$NaBH(OAc)_3$ (6.36 g, 30.0 mmol) was added portionwise to a solution of L-alanine *tert*-butyl ester hydrochloride (1.82 g, 10.0 mmol), benzyl 4-oxopiperidine-1-carboxylate (2.33 g, 10.0 mmol), triethylamine (1.39 mL, 10.0 mmol) in 1,2-dichloroethane (70.0 mL) at 0 ºC (ice bath) under nitrogen in a round bottom flask. Ice bath was removed and after stirring at 27 ºC for 24 h, CH_2Cl_2 (100 mL) was added. The

mixture was washed with saturated $NaHCO_3$ solution 3 times (3 × 100 mL) and then with brine (100 mL). The organic layer was separated, dried over anhydrous $MgSO_4$, and concentrated in vacuo to give the product **16a** 3.48 g (96 % yield) as a colorless oil.

16a (2.90 g, 8.0 mmol) was dissolved in dry dioxane (70.0 mL) in a round bottom flask and heated to 100 ºC under nitrogen atmosphere. Bestmann's ylide (re-crystallized from PhMe, 5.32 g, 17.6 mmol, 2.2 equiv) was added in one portion. Then 10 mL TFA solution (0.4 M in dry dioxane) was added dropwise over 10 mins. The reaction was monitored by ^1H NMR spectroscopy. After completion of reaction (3 h), the solvent was removed under reduced pressure. The resulting oil was dissolved in dichloromethane and filtered through a thin layer of silica gel (~ 1 cm). The silica gel was further washed with 9:1 dichloromethane/ethyl acetate. The solvent was removed under reduced pressure to give the cyclized product contaminated with Ph_3PO.

10 wt % Pd/C (0.85 g, 0.10 equiv Pd) was carefully added to a stirred solution of the crude cyclization product in methanol (60 mL) under nitrogen at 25 ºC. The reaction was evacuated, refilled with N_2, and placed under an atmosphere of H_2 (1 atm, balloon) for 10 h. The reaction mixture was purged with N_2, and filtered over a pad of celite under a gentle vacuum (SAFETY NOTE: Do not let the pad run dry). The celite pad was washed with methanol (2 × 15 mL), and the combined filtrates were concentrated. The residue was purified by flash chromatography (SiO_2, 3 % MeOH/CH_2Cl_2 - 5 % MeOH/CH_2Cl_2 containing 1 % Et_3N) to afford the product **18a** (1.37 g, 68 %) as a white solid.

(*S*)-4-(*tert*-Butoxy)-5-methyl-1-(piperidin-4-yl)-1*H*-pyrrol-2(5*H*)-one (**18a**): white solid, 1.37 g, 68 % yield over two steps; mp = 130.5-131.0 ºC; ^{1}H NMR (300 MHz, $CDCl_3$) d 4.96 (1H, s), 4.01-3.81 (2H, m), 3.18-3.04 (2H, m), 2.71-2.56 (3H, m), 1.83-1.61 (4H, m), 1.39 (9H, s), 1.29 (3H, d, *J* = 6.9 Hz); ^{13}C NMR (75 MHz, $CDCl_3$) d 172.2, 171.8, 95.2, 81.4, 56.6, 49.6, 46.3, 46.2, 32.9, 31.0, 27.4, 18.7. HRMS (ESI+): m/z calcd for $C_{14}H_{24}N_2O_2$ $[M+H]^+$ 253.1916; found 253.1908.

(*S*)-5-Benzyl-4-(*tert*-butoxy)-1-(piperidin-4-yl)-1*H*-pyrrol-2(5*H*)-one (**18f**): white solid, 0.83 g from L-phenylalanine *tert*-butyl ester hydrochloride (1.03 g, 4.00 mmol), 63 % yield over three steps; mp = 218.8-219.6 ºC; ^{1}H NMR (300 MHz, $CDCl_3$) d 7.25-7.13 (5H, m), 5.30-5.22 (1H, br), 4.92 (1H, s), 4.16 (1H, t, *J* = 5.1 Hz), 3.64-3.52 (1H, m), 3.28-3.18 (2H, m), 3.09 (1H, dd, *J* = 5.1, 14.4 Hz), 3.00 (1H, dd, *J* = 5.1, 14.4 Hz), 2.78-2.59 (2H, m), 2.35-2.04 (2H, m), 1.89-1.69 (2H, m), 1.28 (9H, s); ^{13}C NMR (75 MHz, $CDCl_3$) d 174.0, 170.6, 136.0, 129.4, 128.1, 126.7, 96.8, 82.0, 61.6, 50.7, 45.6, 45.5, 36.9, 30.1, 29.7, 27.3. HRMS (ESI+): m/z calcd for $C_{20}H_{28}N_2O_2$ $[M+H]^+$ 329.2229; found 329.2220.

(*S*)-5-((1*H*-Indol-3-yl)methyl)-4-(*tert*-butoxy)-1-(piperidin-4-yl)-1*H*-pyrrol-2(5*H*)-one (**18w**): white solid, 1.12 g from L-tryptophan *tert*-butyl ester hydrochloride (1.48 g, 5.00 mmol), 61 % yield over three steps; mp = 197.9-198.4 ºC; ^{1}H NMR (300 MHz, $CDCl_3$) d 7.58 (1H, d, *J* = 6.6 Hz), 7.32 (1H, d, *J* = 6.6 Hz), 7.16-7.03 (3H, m), 4.83 (1H, s), 4.29 (1H, t, *J* = 4.2 Hz), 3.38-3.19 (2H, m), 3.19-3.04 (2H, m), 2.66-2.53 (2H, m), 2.53-2.32 (1H, br), 2.05-1.69 (4H, m), 1.07 (9H, s); ^{13}C NMR (75 MHz, $CDCl_3$) d 174.0, 170.7, 135.5, 128.1, 122.2, 121.6, 118.8, 118.6, 111.1, 108.7, 96.6, 81.6,

60.7, 50.9, 46.4, 46.3, 32.3, 31.4, 26.8, 25.4. HRMS (ESI+): m/z calcd for $C_{22}H_{29}N_3O_2$ $[M+H]^+$ 368.2338; found 368.2324.

(*S*)-4-(*tert*-Butoxy)-5-(*tert*-butoxymethyl)-1-(piperidin-4-yl)-1*H*-pyrrol-2(5*H*)-one (**3s'**): colorless oil, 0.91 g from *O*-*tert*-butyl-*L*-serine *tert*-butyl ester hydrochloride (1.27 g, 5.00 mmol), 56 % yield over three steps; ^{1}H NMR (300 MHz, $CDCl_3$) d 4.96 (1H, s), 3.97-3.82 (2H, m), 3.67 (1H, dd, *J* = 2.4, 9.3 Hz), 3.39 (1H, dd, *J* = 5.1, 9.3 Hz), 3.18-3.03 (2H, m), 2.64-2.51 (3H, m), 2.08-1.82 (2H, m), 1.79-1.58 (2H, m), 1.38 (9H, s), 1.11 (9H, s); ^{13}C NMR (75 MHz, $CDCl_3$) d 172.9, 168.2, 96.9, 81.4, 72.9, 62.0, 61.2, 50.5, 46.6, 46.4, 32.1, 31.2, 27.4, 27.3. HRMS (ESI+): m/z calcd for $C_{18}H_{32}N_2O_3$ $[M+H]^+$ 325.2491; found 325.2479.

General Procedure for the Syntheses of 19

To a stirred solution of **18a** (126.2mg, 0.5 mmol), **17a** (113.1 mg, 1.0 mmol), and HOAt (149.7 mg, 1.1 mmol) in dry dichloromethane (5 mL) under nitrogen in a round bottom flask, $KHCO_3$ (0.25 g, 2.5 mmol) and EDCI (210.9 mg, 1.1 mmol) was added and the resulting yellow mixture was stirred at room temperature for 10 h. After the reaction is complete, dichloromethane (10 mL) was added to precipitate the urea by-product. The solvent was decanted and the dark red solid was washed twice with dichloromethane (2 × 10 mL). The organic solution was combined and evaporated under reduced pressure. The residue was purified by flash chromatography (SiO_2, 3 % MeOH/CH_2Cl_2 - 5 % MeOH/CH_2Cl_2) to afford the product **19aa** (128.4 mg, 74 % yield) as a colorless oil.

(*S*)-4-(*tert*-Butoxy)-5-methyl-1-(1-((*S*)-2-methyl-5-oxo-2,5-dihydro-1*H*-pyrrol-3-yl)piperidin-4-yl)-1*H*-pyrrol-2(5*H*)-one (**19aa**): colorless oil, 128.4 mg, 74 % yield; ^{1}H NMR (300 MHz, $CDCl_3$) d 6.46 (1H, s), 4.95 (1H, s), 4.66 (1H, s), 4.16 (1H, q, *J* = 6.3 Hz), 4.09-3.95 (1H, m), 3.83 (1H, q, *J* = 6.6 Hz), 3.59-3.40 (2H, m), 3.00-2.83 (2H, m), 1.99-1.66 (4H, m), 1.38 (9H, s), 1.33 (3H, d, *J* = 6.3 Hz), 1.28 (3H, d, *J* = 6.6 Hz); ^{13}C NMR (75 MHz, $CDCl_3$) d 175.9, 172.3, 171.9, 169.4, 95.1, 90.8, 81.6, 56.7, 52.3, 49.2, 48.0, 47.8, 31.2, 28.9, 27.4, 20.5, 18.7. HRMS (ESI+): m/z calcd for $C_{19}H_{29}N_3O_3$ $[M+H]^+$ 348.2287; found 348.2284.

(*S*)-5-Benzyl-4-(*tert*-butoxy)-1-(1-((*S*)-2-methyl-5-oxo-2,5-dihydro-1*H*-pyrrol-3-yl)piperidin-4-yl)-1*H*-pyrrol-2(5*H*)-one (**19af**): colorless oil, 145.9 mg from **18f** (164.0 mg, 0.5 mmol) and **17a** (113.1 mg, 1.0 mmol), 69 % yield; ^{1}H NMR (300 MHz, $CDCl_3$) d 7.31-7.09 (5H, m), 6.17 (1H, s), 4.88 (1H, s), 4.65 (1H, s), 4.16-4.08 (2H, m), 3.62-3.38 (3H, m), 3.11 (1H, dd, *J* = 4.8, 14.7 Hz), 2.92 (1H, dd, *J* = 5.4, 14.7 Hz), 2.85-2.70 (2H, m), 2.20-2.14 (1H, m), 2.02-1.96 (1H, m), 1.83-1.79 (1H, m), 1.72-1.67 (1H, m), 1.38-1.26 (12H, m); ^{13}C NMR (75 MHz, $CDCl_3$) d 175.8, 173.7, 170.3, 169.4, 136.2, 129.2, 128.2, 126.9, 96.8, 90.8, 81.9, 61.7, 52.2, 51.1, 48.1, 47.8, 37.3, 29.8, 29.1, 27.3, 20.5. HRMS (ESI+): m/z calcd for $C_{25}H_{33}N_3O_3$ $[M+H]^+$ 424.2600; found 424.2605.

(*S*)-5-((1*H*-Indol-3-yl)methyl)-4-(*tert*-butoxy)-1-(1-((*S*)-2-methyl-5-oxo-2,5-dihydro-1*H*-pyrrol-3-yl)piperidin-4-yl)-1*H*-pyrrol-2(5*H*)-one (**19aw**): light yellow oil, 145.5 mg from **18w** (183.5 mg, 0.5 mmol) and **17a** (113.1 mg, 1.0 mmol), 63 % yield; ^{1}H NMR (300 MHz, $CDCl_3$) d 9.05 (1H, s), 7.53 (1H, d, *J* = 7.8 Hz), 7.34 (1H, d, *J* = 7.5 Hz), 7.16-7.04 (2H, m), 7.01 (1H, d, *J* = 2.1 Hz), 6.19 (1H, s), 4.91 (1H, s), 4.60 (1H, s),

4.26 (1H, t, J = 5.1 Hz), 4.00 (1H, q, J = 6.3 Hz), 3.89-3.77 (1H, m), 3.44-3.06 (4H, m), 2.83-2.64 (2H, m), 2.12-1.54 (4H, m), 1.28 (3H, d, J = 6.3 Hz), 1.23 (9H, s); ^{13}C NMR (75 MHz, $CDCl_3$) d 176.1, 173.8, 170.9, 169.5, 135.8, 127.7, 122.2, 121.8, 119.0, 118.4, 111.4, 109.3, 96.4, 90.4, 81.9, 61.3, 52.2, 50.4, 48.1, 47.8, 30.3, 28.7, 27.1, 26.6, 20.4. HRMS (ESI+): m/z calcd for $C_{27}H_{34}N_4O_3$ $[M+H]^+$ 463.2709; found 463.2692.

(*S*)-4-(*tert*-Butoxy)-5-(*tert*-butoxymethyl)-1-(1-((*S*)-2-methyl-5-oxo-2,5-dihydro-1*H*-pyrrol-3-yl)piperidin-4-yl)-1*H*-pyrrol-2(5*H*)-one (**19as'**): colorless oil, 100.4 mg from **18s'** (130.0 mg, 0.4 mmol) and **17a** (90.5 mg, 0.8 mmol), 60 % yield; ^{1}H NMR (300 MHz, $CDCl_3$) d 6.14 (1H, s), 4.99 (1H, s), 4.66 (1H, s), 4.15 (1H, q, J = 6.6 Hz), 4.10-3.95 (1H, m), 3.92 (1H, dd, J = 2.1, 6.6 Hz), 3.66 (1H, dd, J = 2.4, 9.6 Hz), 3.59-3.39 (2H, m), 3.28 (1H, dd, J = 6.9, 9.6 Hz), 3.00-2.81 (2H, m), 2.34-2.00 (2H, m), 1.82-1.59 (2H, m), 1.41 (9H, s), 1.34 (3H, d, J = 6.6 Hz), 1.14 (9H, s); ^{13}C NMR (75 MHz, $CDCl_3$) d 176.1, 172.8, 169.6, 168.1, 96.7, 90.2, 81.8, 73.4, 62.7, 62.3, 52.3, 49.9, 48.2, 47.9, 30.2, 28.7, 27.5, 27.4, 20.5. HRMS (ESI+): m/z calcd for $C_{23}H_{37}N_3O_4$ $[M+H]^+$ 420.2862; found 420.2878.

(*S*)-1-(1-((*S*)-2-Benzyl-5-oxo-2,5-dihydro-1*H*-pyrrol-3-yl)piperidin-4-yl)-4-(*tert*-butoxy)-5-(*tert*-butoxymethyl)-1*H*-pyrrol-2(5*H*)-one (**19fs'**): colorless oil, 45.5 mg from **18s'** (130.0 mg, 0.4 mmol) and **17f** (151.4 mg, 0.8 mmol), 23 % yield; ^{1}H NMR (300 MHz, $CDCl_3$) d 7.33-7.18 (5H, m), 5.22 (1H, s), 5.01 (1H, s), 4.73 (1H, d, J = 2.4 Hz), 4.28 (1H, dd, J = 4.5, 9.9 Hz), 4.13-3.99 (1H, m), 3.95 (1H, dd, J = 2.4, 6.9 Hz), 3.73-3.46 (3H, m), 3.34-3.27 (1H, m), 3.16 (1H, dd, J = 2.7, 7.8 Hz), 3.09-2.89 (2H, m), 2.55 (1H, dd, J = 9.6, 13.8 Hz), 2.21-2.09 (2H, m), 1.88-1.67 (2H, m), 1.42 (9H, s), 1.14 (9H,

s); ^{13}C NMR (75 MHz, $CDCl_3$) d 175.2, 172.8, 168.1, 167.6, 137.1, 129.0, 128.8, 127.1, 96.8, 91.9, 81.8, 73.4, 62.8, 62.4, 57.7, 50.0, 48.5, 48.1, 41.3, 30.2, 28.7, 27.5, 27.4. HRMS (ESI+): m/z calcd for $C_{29}H_{41}N_3O_4$ $[M+H]^+$ 496.3176; found 496.3163.

(*S*)-4-(*tert*-Butoxy)-5-methyl-1-(1-((*S*)-2-(2-(methylthio)ethyl)-5-oxo-2,5-dihydro-1*H*-pyrrol-3-yl)piperidin-4-yl)-1*H*-pyrrol-2(5*H*)-one (**19ma**): light yellow oil, 50.9 mg from **18a** (126.2mg, 0.5 mmol) and **17m** (173.2 mg, 1.0 mmol), 25 % yield; ^{1}H NMR (300 MHz, $CDCl_3$) d 6.45 (1H, s), 5.00 (1H, s), 4.78 (1H, d, *J* = 1.5 Hz), 4.33-4.29 (2H, m), 4.11-3.98 (1H, m), 3.87 (1H, q, *J* = 6.6 Hz), 3.60-3.42 (2H, m), 3.04-2.91 (2H, m), 2.60 (2H, t, *J* = 7.8 Hz), 2.12 (3H, s), 2.10-1.71 (6H, m), 1.44 (9H, s), 1.32 (3H, d, *J* = 6.6 Hz); ^{13}C NMR (75 MHz, $CDCl_3$) d 176.0, 172.4, 171.9, 167.7, 95.2, 92.4, 81.7, 56.7, 55.6, 49.2, 48.3, 48.1, 33.8, 31.2, 30.4, 28.9, 27.4, 18.7, 15.7. HRMS (ESI+): m/z calcd for $C_{21}H_{33}N_3O_3S$ $[M+H]^+$ 408.2321; found 408.2314.

Procedure for the Syntheses of 15aaa

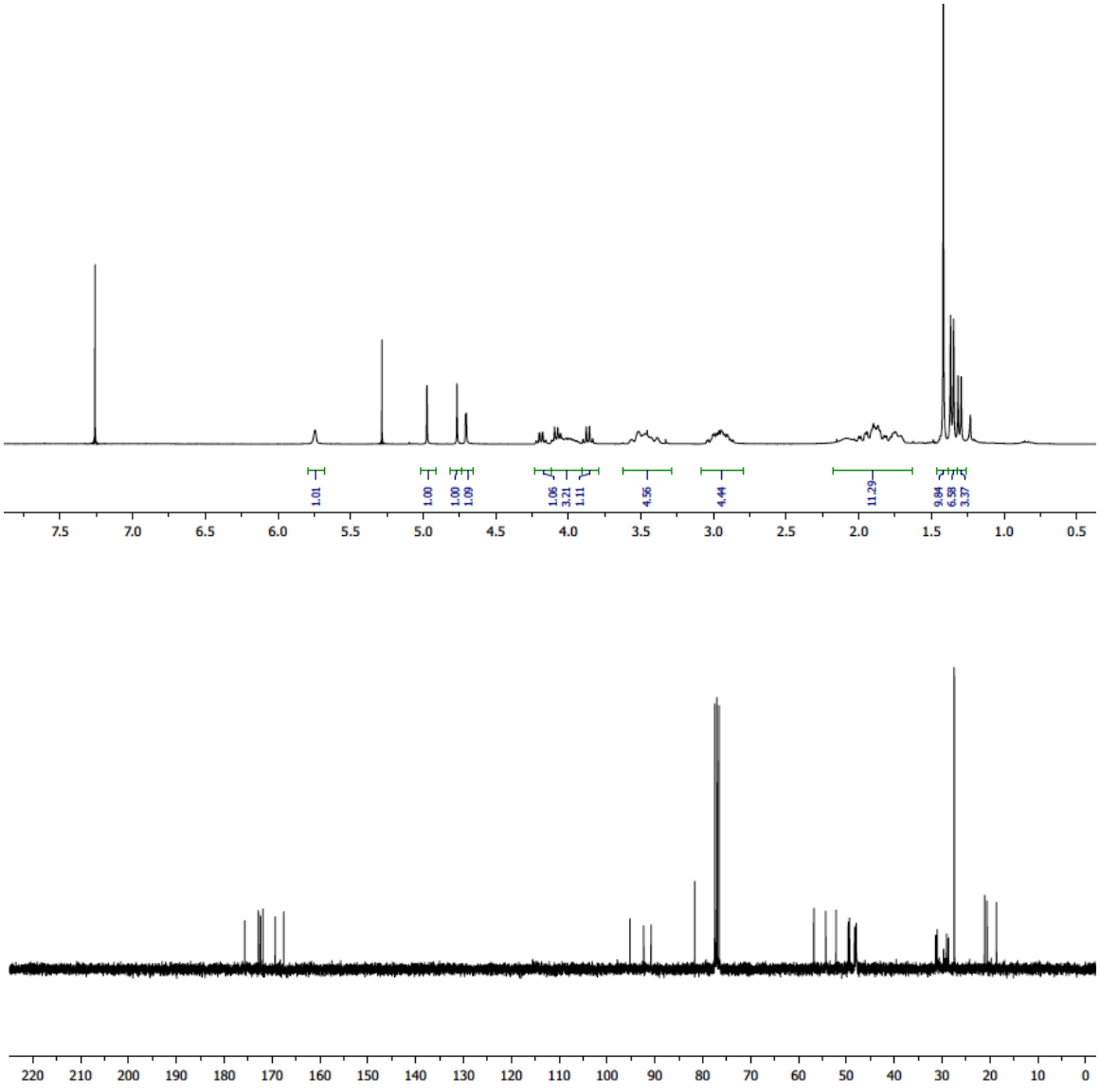
19aa (277.7 mg, 0.5 mmol) was treated with TFA/CH_2Cl_2 (1:1, 5 mL) at 25 ºC. The reaction was stirred in a vented teflon-capped flask until complete disappearance of starting material (monitored by NMR spectroscopy, ~ 3 h). Toluene (10 mL) was added and the reaction mixture was concentrated *in vacuo*. Toluene (2 × 10 mL) was used to azeotrope residual TFA. The crude oil was dissolved in dichloromethane (4 mL) and filtered through a thin layer of silica gel (~ 0.5 cm) to remove some polar impurities. The silica gel was washed with 5 % MeOH/CH_2Cl_2 5 times (5 × 5 mL). The organic fractions were combined and the solvent was removed in vacuo to give the deprotected

19aa as a white solid. The white solid was dried under high vacuum and used directly in the next step without further purification.

To a stirred solution of **18a** (75.6 mg, 0.3 mmol), deprotected **19aa** from previous step, and HOAt (89.8 mg, 0.66 mmol) in dry dichloromethane (2 mL) under nitrogen, $KHCO_3$ (150 mg, 1.5 mmol) and EDCI (126.5 mg, 0.66 mmol) was added and the resulting yellow mixture was stirred at room temperature for 16 h. After the reaction was complete, dichloromethane (4 mL) was added to precipitate the urea by-product. The solvent was decanted and the light red solid was washed twice with dichloromethane (2 × 3 mL). The organic solution was combined and evaporated under reduced pressure. The residue was purified by flash chromatography (SiO_2, 4 % MeOH/CH_2Cl_2 - 6 % MeOH/CH_2Cl_2) to afford the product **15aaa**.

(*S*)-4-(*tert*-Butoxy)-5-methyl-1-(1-((*S*)-2-methyl-1-(1-((*S*)-2-methyl-5-oxo-2,5-dihydro-1H-pyrrol-3-yl)piperidin-4-yl)-5-oxo-2,5-dihydro-1H-pyrrol-3-yl)piperidin-4-yl)-1*H*-pyrrol-2(5*H*)-one (**15aaa**): colorless oil, 64.6 mg, 41 % yield over two steps; ^{1}H NMR (300 MHz, $CDCl_3$) d 5.74 (1H, s), 4.97 (1H, s), 4.77 (1H, s), 4.71 (1H, d, *J* = 1.8 Hz), 4.18 (1H, q, *J* = 6.3 Hz), 4.12-3.91 (3H, m), 3.86 (1H, q, *J* = 6.3 Hz), 3.59-3.33 (4H, m), 3.06-2.84 (4H, m), 2.19-1.68 (8H, m), 1.42 (9H, s), 1.35 (6H, d, *J* = 6.3 Hz), 1.30 (3H, d, *J* = 6.6 Hz); ^{13}C NMR (75 MHz, $CDCl_3$) d 175.7, 172.8, 172.4, 171.9, 169.3, 167.6, 95.2, 92.3, 90.8, 81.7, 56.8, 54.3, 52.1, 49.6, 49.3, 48.3, 48.1, 48.0, 47.9, 31.3, 31.0, 29.0, 28.7, 27.4, 21.1, 20.6, 18.6. HRMS (MALDI+): m/z calcd for $C_{29}H_{43}N_5O_4$ $[M+H]^+$ 526.3388; found 526.3384.

(*S*)-4-(*tert*-Butoxy)-5-methyl-1-(1-((*S*)-2-methyl-1-(1-((*S*)-2-methyl-5-oxo-2,5-dihydro-1H-pyrrol-3-yl)piperidin-4-yl)-5-oxo-2,5-dihydro-1H-pyrrol-3-yl)piperidin-4-yl)-1*H*-pyrrol-2(5*H*)-one (**15aaa**):



B. Synthesis Of β-Ketoesters And β-Ketothioesters

Synthesis Of β-Ketoesters And β-Ketothioesters With 2,2,6-Trimethyl-4*H*-1,3-dioxin-4-one

A solution of 20 mmol phenol or thiol and 20 mmol 2,2,6-trimethyl-4*H*-1,3-dioxin-4-one in 10 mL xylenes was placed in a 50 mL Erlenmeyer flask. The the solution was vigorously stirred at 145 °C. After 30 minutes, the reaction mixture was cooled and the xylene was removed. The residue was purified by flash chromatography (hexanes/CH_2Cl_2 3:1 to 1:2) to give the desired product.[10]

Procedure For The Alkylation Of β-Ketothioesters

NaH (132 mg, 5.5 mmol) was added to a solution of *S*-butyl 3-oxobutanethioate (872 mg, 5 mmol) in 20 mL anhydrous 1,2-dimethoxylethane (DME) at 0 °C under N_2. After stirring the resulting mixture at 0 °C for 30 min, anhydrous allyl bromide (480 uL, 5.5 mmol) was added dropwise over 5 min. The mixture was stirred at room temperature for 18 h. The reaction was quenched with saturated $NaHCO_3$ (50 mL) and the aqueous phase was extracted with DCM three times (50 mL × 3). The combined organic phase was washed with brine and dried over $MgSO_4$. The solvent was removed under vacuum and the crude product was purified by flash chromatography (hexanes/CH_2Cl_2 5:1 to 1:1).[11]

C. Synthesis Of β-Enamino Esters And β-Enamino Thioesters

General Procedures

Method A with EDCI and HOAt:

EDCI (144 mg, 0.75 mmol) and HOAt (82 mg, 0.6 mmol) were added to a mixture of the β-keto substrate (0.5 mmol) and $KHCO_3$ powder (500 mg, 10 equiv.) in 2.5 mL $CHCl_3$ under N_2 at 25 °C. The mixture was stirred for 10 min, followed by the addition of amine nucleophile. The resulting light yellow mixture was stirred at 25 °C for 12 h. 10 mL of DCM was added and the mixture was filtered through Celite. Solvent was removed under vacuum to give the crude product. The product was dissolved in DCM and filtered through a pad of silica gel and the silica gel was washed with DCM a couple times to give the pure β-enamino ester or β-enamino thioester. For the synthesis of β-enamino amides, the crude product was purified by flash chromatography (3 % - 6 % MeOH in CH_2Cl_2). This method was applied to the syntheses of compounds **25**.

Method B with PyBOP and imidazole:

PyBOP (312 mg, 0.6 mmol) and imidazole (170 mg, 2.5 mmol) were added to a mixture of the β-keto substrate (0.5 mmol) in 2.5 mL $CHCl_3$ under N_2 at 25 °C. The mixture was stirred for 10 min, followed by the addition of amine nucleophile. The resulting mixture was stirred at 25 °C for 12 h. 10 mL of DCM was added and the mixture was filtered through a short layer of silica gel and the silica gel was washed with DCM a couple times to give the pure β-enamino ester or β-enamino thioester unless

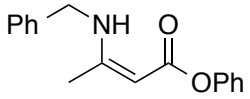
otherwise noted for specific compounds. This method was applied to the syntheses of compounds **20**.

Method C with PyBOP, HOAt and $KHCO_3$:

Same as method B except that HOAt (82 mg, 0.6 mmol) and $KHCO_3$ powder (500 mg, 10 equiv.) were used instead of imidazole. This method was applied to the syntheses of compounds **21**.

Compound **20a**

Phenyl (*Z*)-3-(benzylamino)but-2-enoate

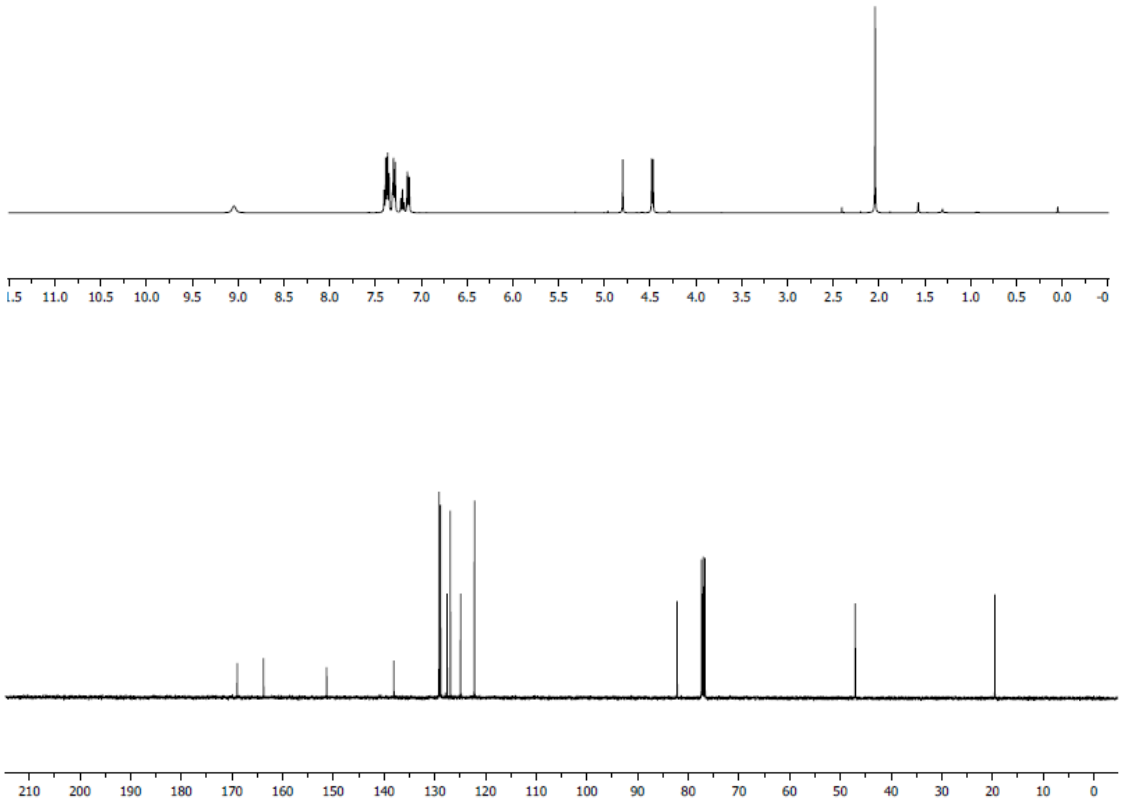


White solid, 87 %; mp = 116.8-118.0 °C;

^{1}H-NMR (400 MHz, $CDCl_3$) δ 9.04 (s, 1H), 7.46-7.24 (m, 7H), 7.24-7.09 (m, 3H), 4.80 (s, 1H), 4.47 (d, *J* = 6.2 Hz, 2H), 2.04 (s, 3H);

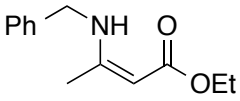
^{13}C-NMR (100 MHz, $CDCl_3$) δ 168.9, 163.7, 151.3, 138.0, 129.1, 128.8, 127.5, 126.9, 124.9, 122.1, 82.1, 47.0, 19.5;

HRMS (ESI) m/z calcd for $C_{17}H_{17}NO_2$ $(M+H)^+$ 268.1338; found 268.1325.



Compound **20b**

Ethyl (*Z*)-3-(benzylamino)but-2-enoate

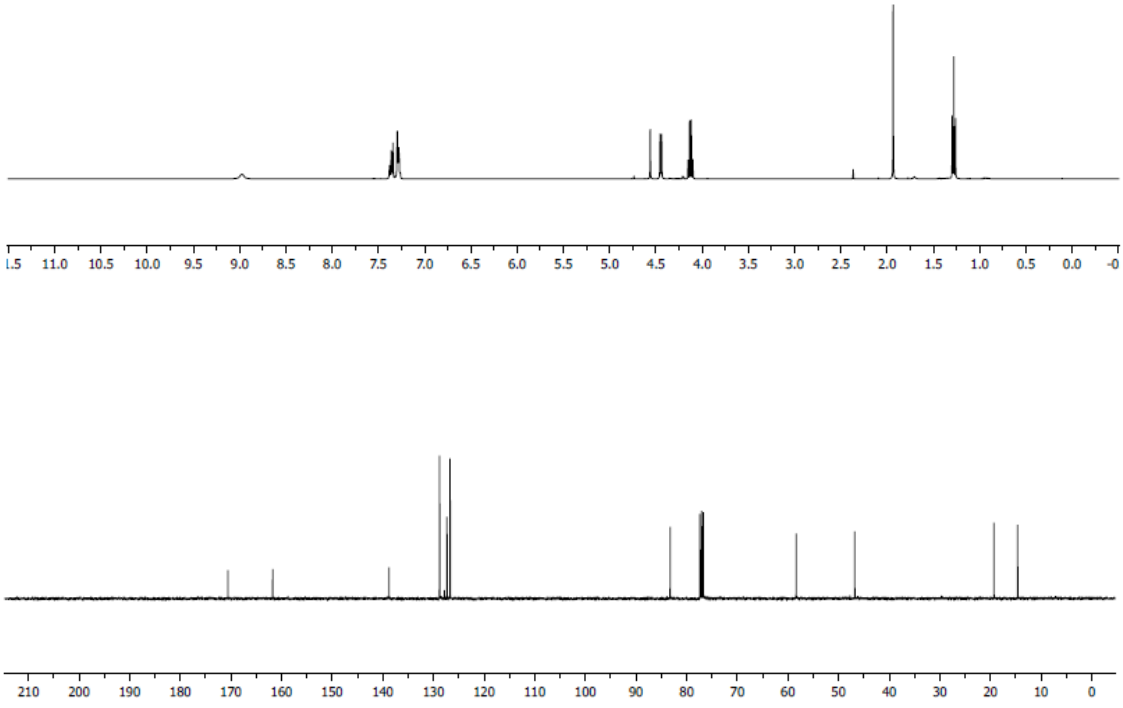


Colorless oil, 91 %;

^{1}H-NMR (400 MHz, $CDCl_3$) δ 8.98 (s, 1H), 7.41-7.23 (m, 5H), 4.56 (s, 1H), 4.45 (d, *J* = 6.4 Hz, 2H), 4.13 (q, *J* = 7.1 Hz, 2H), 1.94 (s, 3H), 1.28 (t, *J* = 7.1 Hz, 3H);

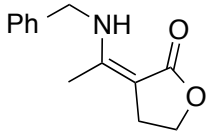
^{13}C-NMR (100 MHz, $CDCl_3$) δ 170.5, 161.7, 138.7, 128.7, 127.3, 126.7, 83.2, 58.3, 46.8, 19.3, 14.6;

MS (ESI+) m/z calcd for $C_{13}H_{17}NO_2$ $(M+H)^+$ 220.13; found 220.12.



Compound **20c**

(*Z*)-3-(1-(Benzylamino)ethylidene)dihydrofuran-2(3*H*)-one

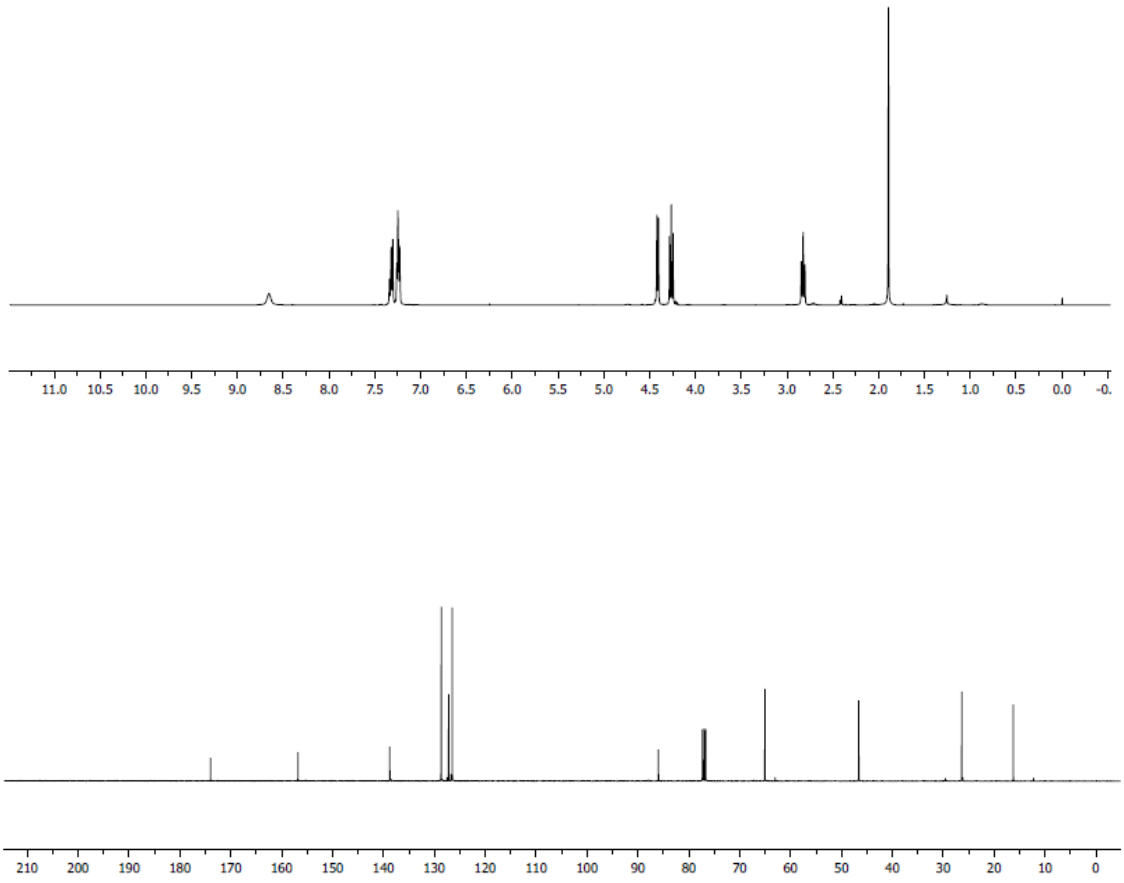


White solid, 92 %; mp = 104.6-106.2 °C;

^{1}H-NMR (400 MHz, $CDCl_3$) δ 8.66 (s, 1H), 7.38-7.18 (m, 5H), 4.41 (d, J = 6.5 Hz, 2H), 4.27 (t, J = 7.9 Hz, 2H), 2.82 (t, J = 7.9 Hz, 2H), 1.89 (s, 3H);

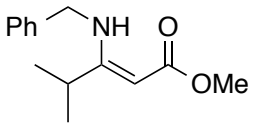
^{13}C-NMR (100 MHz, $CDCl_3$) δ 173.9, 156.7, 138.7, 128.6, 127.2, 126.4, 85.9, 65.0, 46.6, 26.3, 16.2;

HRMS (ESI) m/z calcd for $C_{13}H_{15}NO_2$ $(M+H)^+$ 218.1181; found 218.1175.



Compound **20d**

Methyl (*Z*)-3-(benzylamino)-4-methylpent-2-enoate

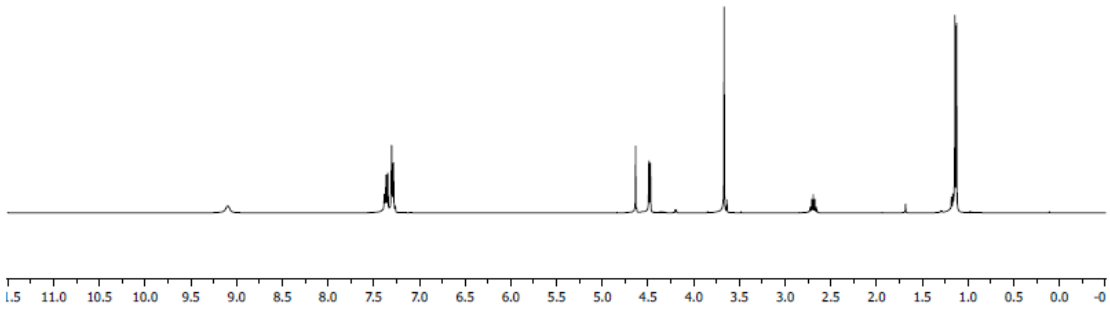


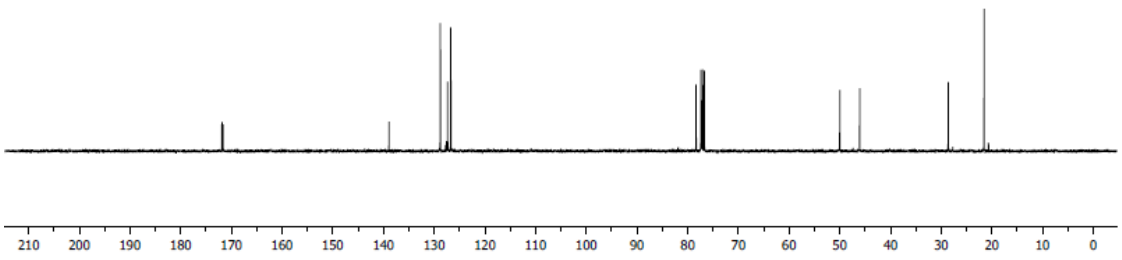
Colorless oil, 88 %;

^{1}H-NMR (400 MHz, $CDCl_3$) δ 9.10 (s, 1H), 7.49-7.22 (m, 5H), 4.63 (s, 1H), 4.48 (d, J = 6.3 Hz, 2H), 3.67 (s, 3H), 2.75-2.65 (m, 1H), 1.13 (d, J = 6.9 Hz, 6H);

^{13}C-NMR (100 MHz, $CDCl_3$) δ 171.8, 171.6, 138.9, 128.7, 127.3, 126.7, 78.3, 49.9, 46.0, 28.6, 21.5;

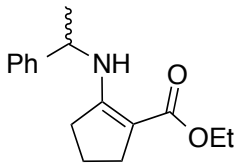
HRMS (ESI) m/z calcd for $C_{14}H_{19}NO_2$ $(M+H)^+$ 234.1494; found 234.1483.





Compound **20e**

Ethyl 2-((1-phenylethyl)amino)cyclopent-1-ene-1-carboxylate

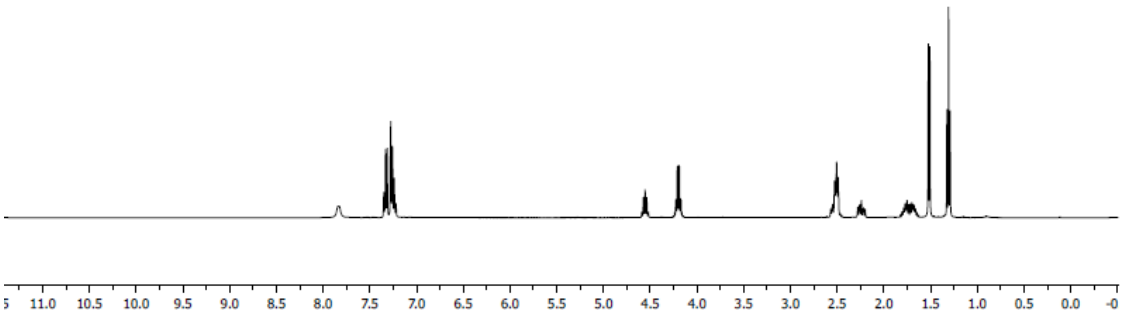


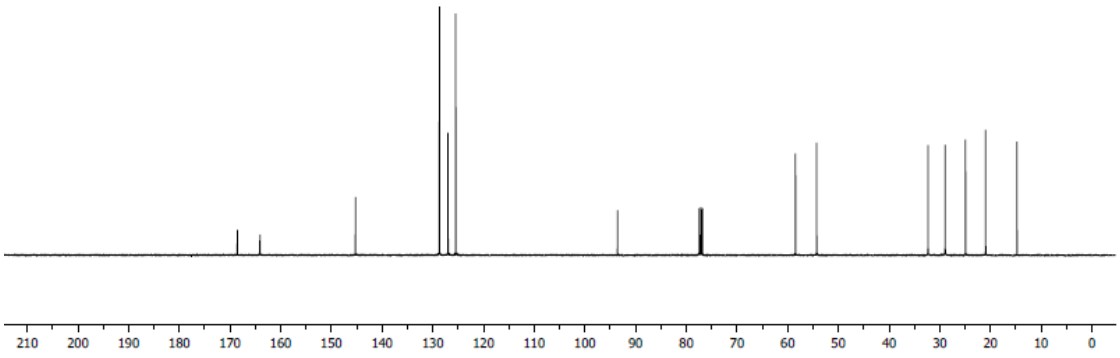
Colorless oil, 89 %;

^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.83 (s, 1H), 7.38-7.19 (m, 5H), 4.61-4.51 (m, 1H), 4.20 (q, *J* = 7.1, 2H), 2.58-2.45 (m, 3H), 2.32-2.19 (m, 1H), 1.80-1.66 (m, 2H), 1.52 (d, *J* = 6.8 Hz, 3H), 1.31 (t, *J* = 7.1 Hz, 3H);

^{13}C-NMR (100 MHz, $CDCl_3$) δ 168.5, 164.0, 145.2, 128.6, 127.0, 125.4, 93.5, 58.4, 54.2, 32.3, 28.9, 24.9, 20.9, 14.7;

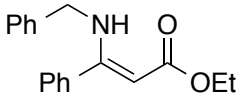
HRMS (ESI) m/z calcd for $C_{16}H_{21}NO_2$ $(M+H)^+$ 260.1651; found 260.1639.





Compound **20f**

Ethyl (*Z*)-3-(benzylamino)-3-phenylacrylate

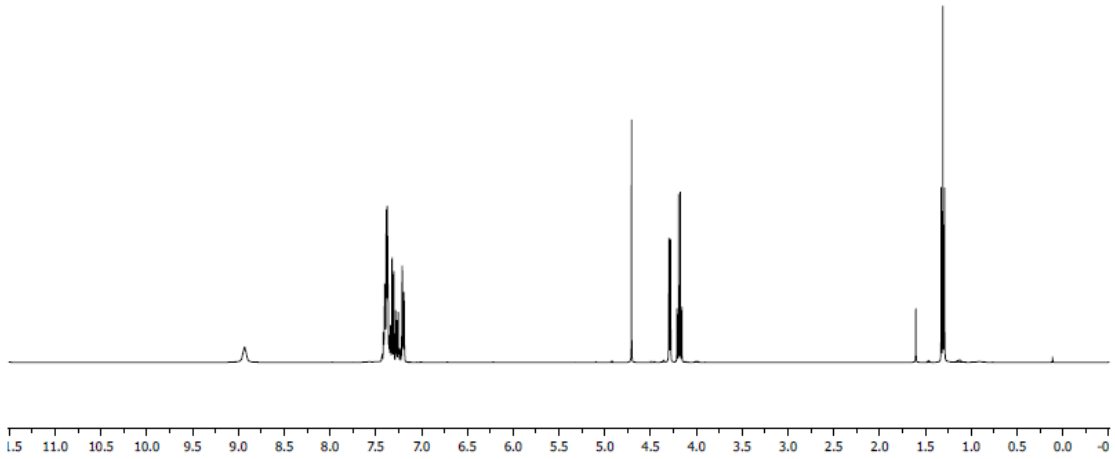


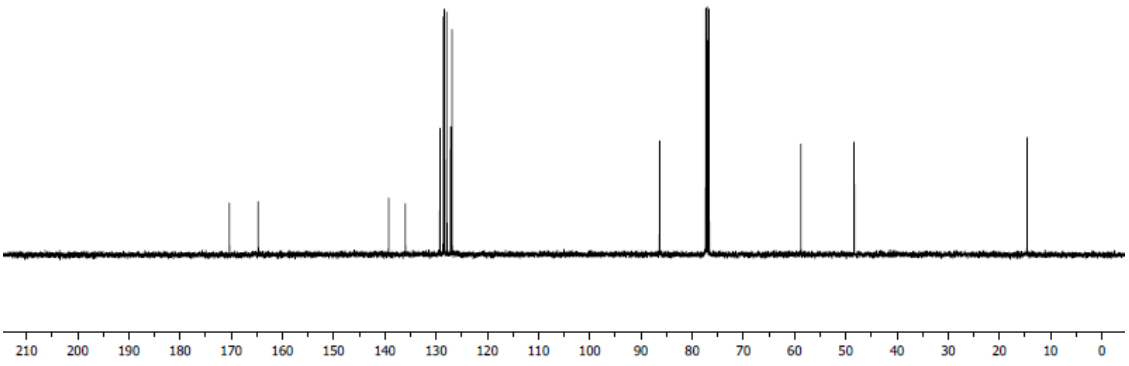
White solid, 75 % (this compound was purified by flash chromatography to remove unreacted β-keto ester starting material); mp = 62.5-64.4 °C;

^{1}H-NMR (400 MHz, $CDCl_3$) δ 8.93 (s, 1H), 7.44-7.17 (m, 10H), 4.71 (s, 1H), 4.29 (d, *J* = 6.5 Hz, 2H), 4.18 (q, *J* = 7.1 Hz, 2H), 1.31 (t, *J* = 7.1 Hz, 3H);

^{13}C-NMR (100 MHz, $CDCl_3$) δ 170.3, 164.7, 139.2, 135.9, 129.2, 128.6, 128.3, 127.8, 127.1, 126.8, 86.3, 58.7, 48.3, 14.5;

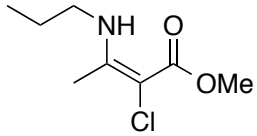
HRMS (ESI) m/z calcd for $C_{18}H_{19}NO_2$ $(M+H)^+$ 282.1494; found 282.1482.





Compound **20g**

Methyl (*E*)-2-chloro-3-(propylamino)but-2-enoate

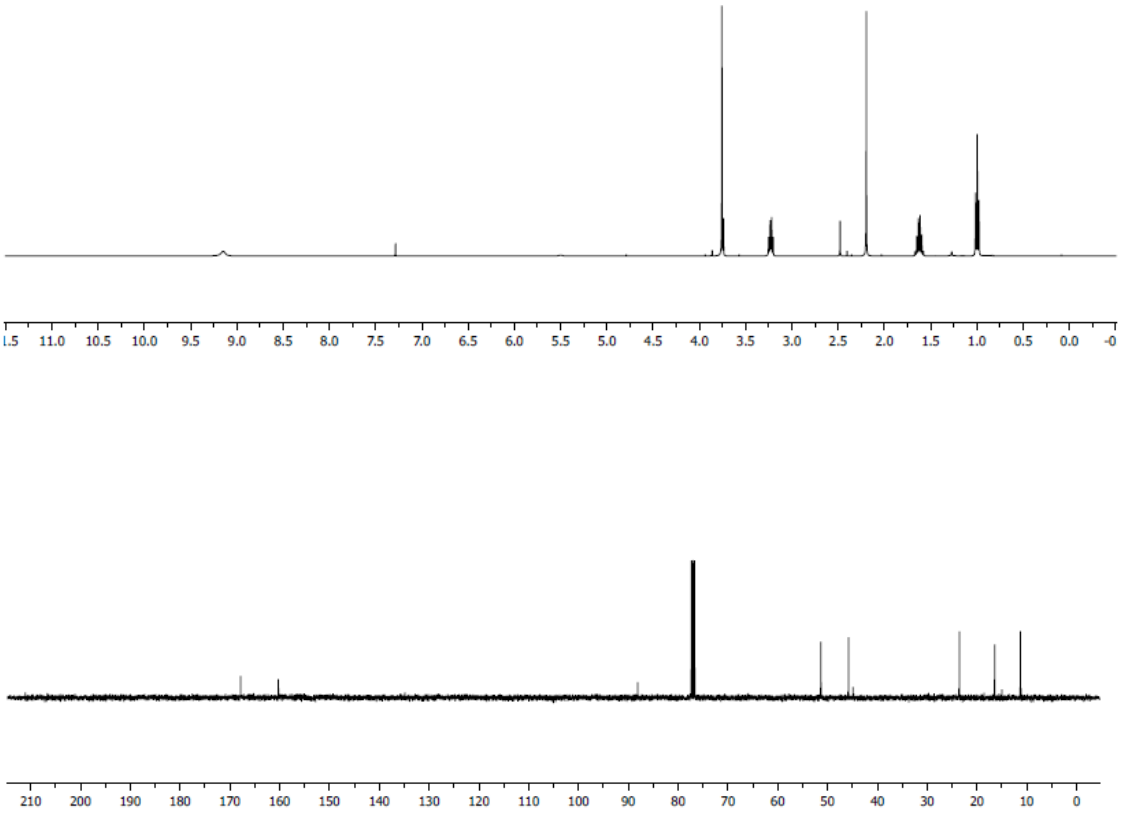


Light yellow oil, 90 %;

^{1}H-NMR (400 MHz, $CDCl_3$) δ 9.15 (s, 1H), 3.77 (s, 3H), 3.23 (td, J = 7.0, 5.7 Hz, 2H), 2.19 (s, 3H), 1.8-1.57 (m, 2H), 1.00 (t, J = 7.4 Hz, 3H);

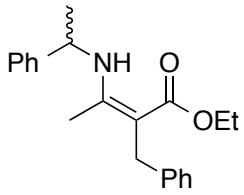
^{13}C-NMR (100 MHz, $CDCl_3$) δ 167.8, 160.2, 88.1, 51.3, 45.7, 23.5, 16.4, 11.2;

HRMS (ESI+) m/z calcd for $C_8H_{14}ClNO_2$ $(M+H)^+$ 192.0791; found 192.0786.



Compound **20h**

Ethyl (*Z*)-2-benzyl-3-((1-phenylethyl)amino)but-2-enoate

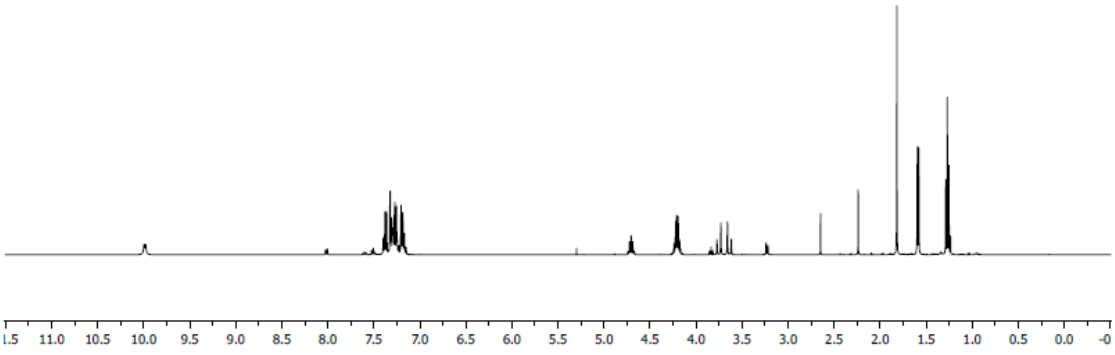


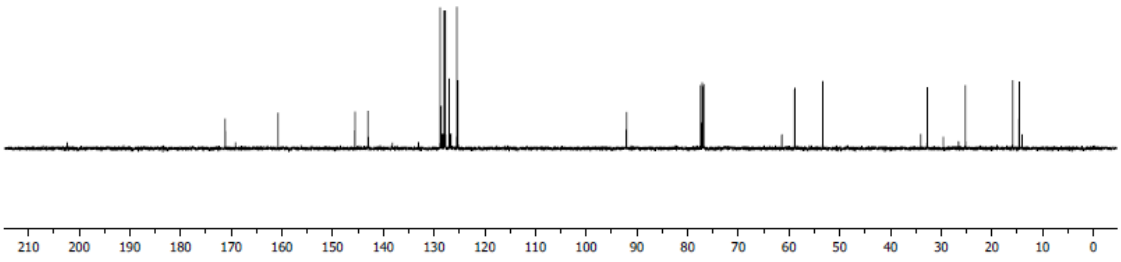
Colorless oil, 71 %;

^{1}H-NMR (400 MHz, $CDCl_3$) δ 9.99 (d, J = 7.2 Hz, 1H), 7.43-7.12 (m, 10H), 4.75-4.62 (m, 1H), 4.31-4.12 (m, 2H), 3.75 (d, J = 16.9 Hz, 1H), 3.64 (d, J = 16.9 Hz, 1H), 1.82 (s, 3H), 1.59 (d, J = 6.8 Hz, 3H), 1.27 (t, J = 7.1 Hz, 3H);

^{13}C-NMR (100 MHz, $CDCl_3$) δ 171.1, 160.7, 145.6, 143.0, 128.7, 128.0, 127.7, 126.9, 125.5, 125.3, 92.0, 58.8, 53.3, 32.7, 25.2, 15.8, 14.6;

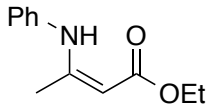
HRMS (ESI+) m/z calcd for $C_{21}H_{25}NO_2$ $(M+H)^+$ 324.1964; found 324.1971.





Compound **20i**

Ethyl (*Z*)-3-(phenylamino)but-2-enoate

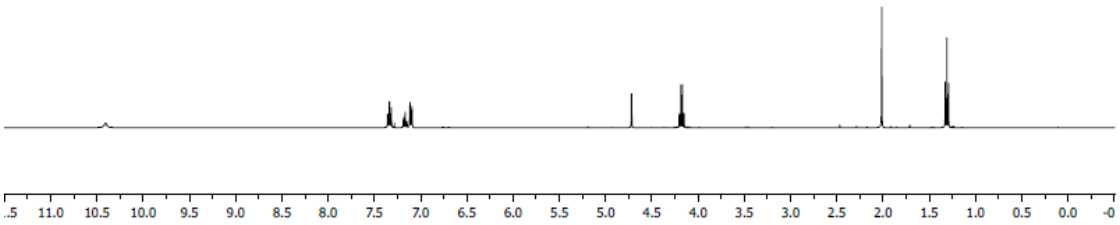


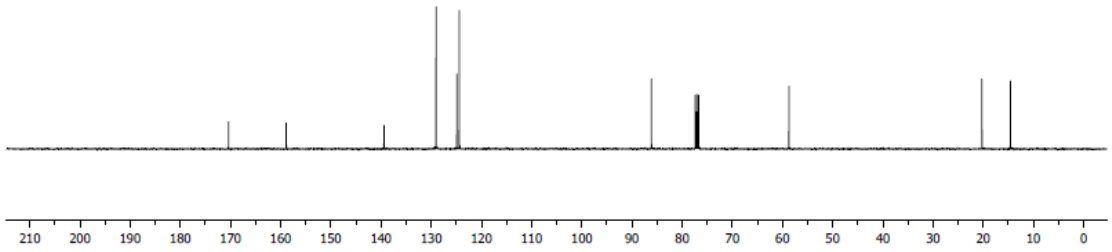
Colorless oil, 82 %;

^{1}H-NMR (400 MHz, $CDCl_3$) δ 10.41 (s, 1H), 7.40-7.26 (m, 2H), 7.21-7.07 (m, 3H), 4.72 (d, *J* = 0.5 Hz, 1H), 4.18 (q, *J* = 7.1 Hz, 2H), 2.03 (s, 3H), 1.31 (t, *J* = 7.1 Hz, 3H);

^{13}C-NMR (100 MHz, $CDCl_3$) δ 170.3, 158.8, 139.4, 129.0, 124.9, 124.4, 86.1, 58.7, 20.2, 14.5;

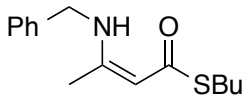
HRMS (ESI+) m/z calcd for $C_{12}H_{15}NO_2$ $(M+H)^+$ 206.1181; found 206.1177.


.5 11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0


210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0

Compound **21a**

S-Butyl (*Z*)-3-(benzylamino)but-2-enethioate

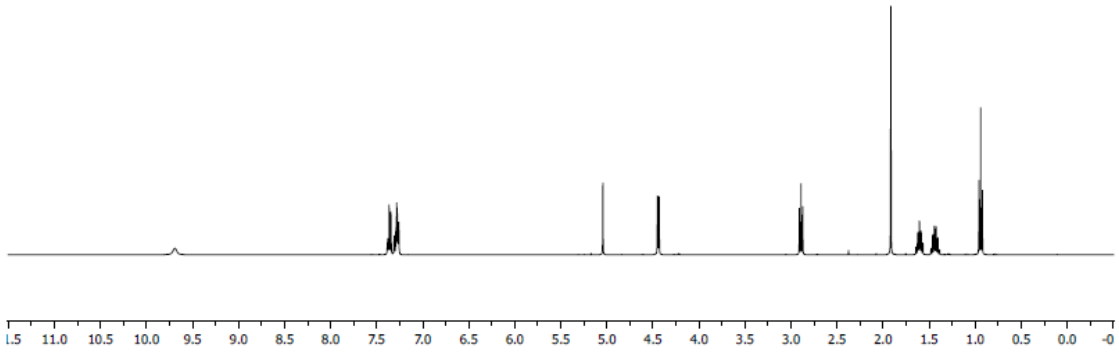


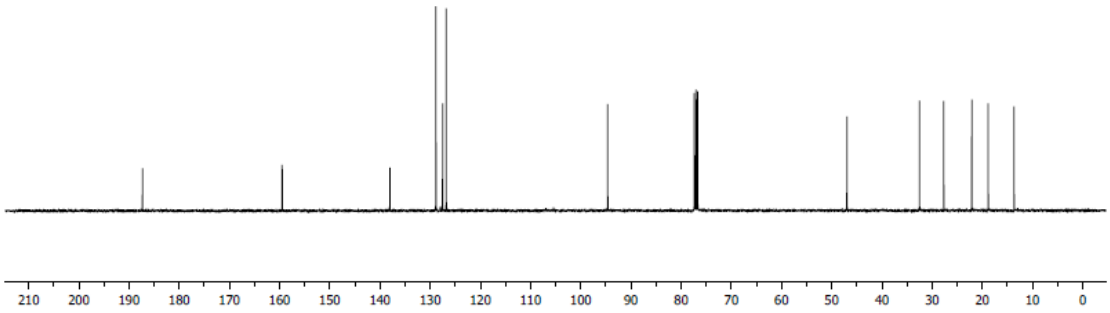
Colorless oil, 90 %;

^{1}H-NMR (400 MHz, $CDCl_3$) δ 9.69 (s, 1H), 7.42-7.23 (m, 5H), 5.04 (s, 1H), 4.44 (d, *J* = 6.4 Hz, 2H), 2.89 (t, *J* = 7.3 Hz, 2H), 1.92 (s, 3H), 1.65-1.57 (m, 2H), 1.50-1.38 (m, 2H), 0.94 (t, *J* = 7.3 Hz, 3H);

^{13}C-NMR (100 MHz, $CDCl_3$) δ 187.3, 159.4, 138.0, 128.8, 127.5, 126.8, 94.6, 46.9, 32.5, 27.6, 22.0, 18.8, 13.6;

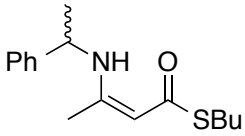
HRMS (ESI) m/z calcd for $C_{15}H_{21}NOS$ $(M+H)^+$ 264.1422; found 264.1410.





Compound **21b**

S-Butyl (*Z*)-3-((1-phenylethyl)amino)but-2-enethioate

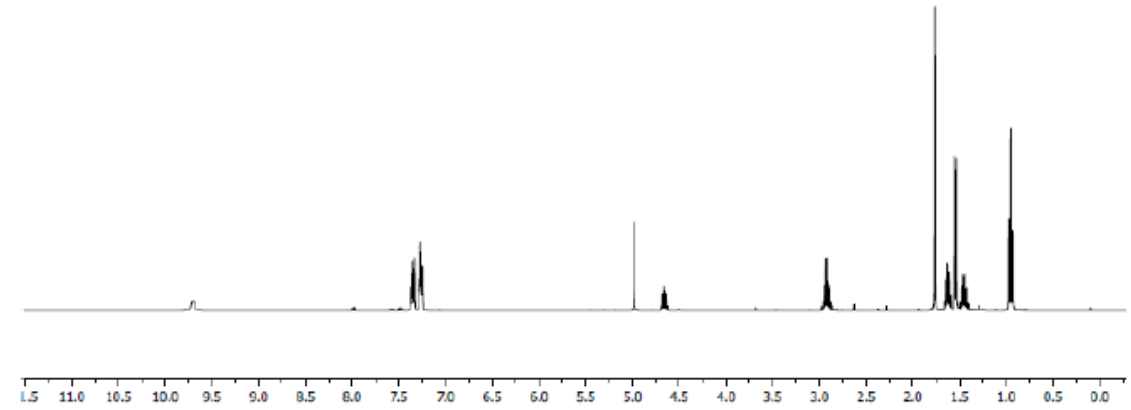


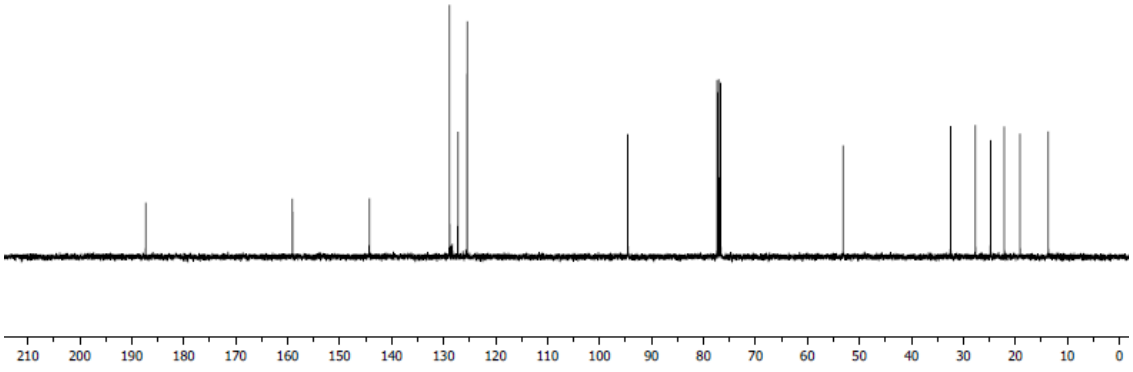
Light yellow oil, 94 %;

^{1}H-NMR (400 MHz, $CDCl_3$) δ 9.70 (s, 1H), 7.41-7.21 (m, 5H), 4.98 (s, 1H), 4.71-4.62 (m, 1H), 2.99-2.86 (m, 2H), 1.77 (s, 3H), 1.68-1.59 (m, 2H), 1.54 (d, J = 6.8 Hz, 3H), 1.49-1.40 (s, 2H), 0.95 (t, J = 7.3 Hz, 3H);

^{13}C-NMR (100 MHz, $CDCl_3$) δ 187.3, 159.0, 144.2, 128.8, 127.2, 125.4, 94.5, 53.1, 32.4, 27.6, 24.7, 22.1, 19.1, 13.6;

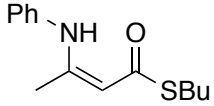
HRMS (ESI) m/z calcd for $C_{16}H_{23}NOS$ $(M+H)^+$ 278.1579; found 278.1573.





Compound **21c**

S-Butyl (*Z*)-3-(phenylamino)but-2-enethioate

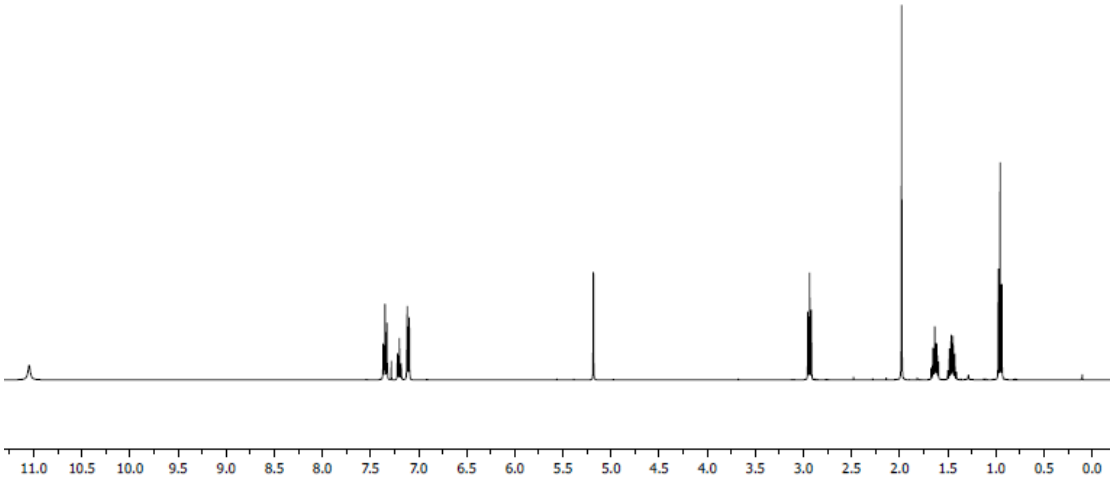


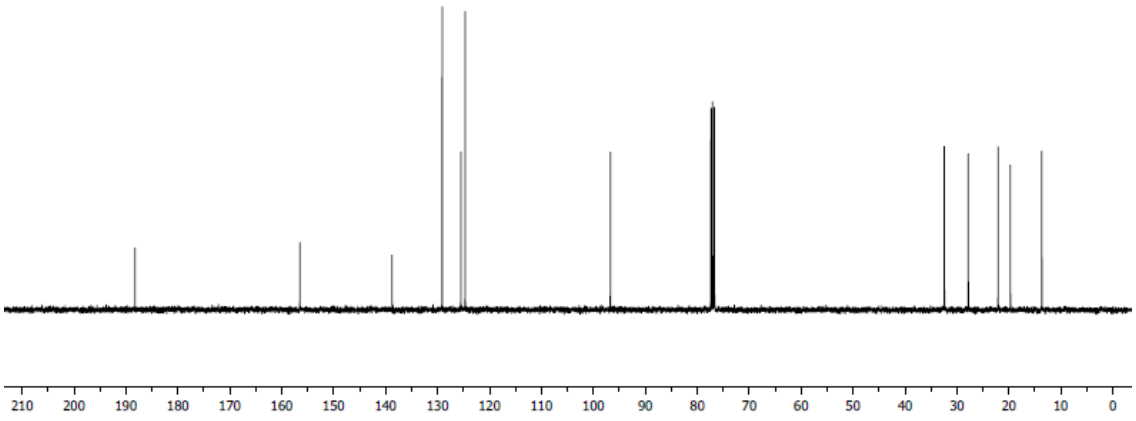
Colorless oil, 92 %;

^{1}H-NMR (400 MHz, $CDCl_3$) δ 11.05 (s, 1H), 7.41-7.26 (m, 2H), 7.20 (t, *J* = 7.4 Hz, 1H), 7.11 (d, *J* = 7.5 Hz, 2H), 5.18 (s, 1H), 2.94 (t, *J* = 7.3 Hz, 2H), 1.98 (s, 3H), 1.70-1.57 (m, 2H), 1.49-1.40 (m, 2H), 0.96 (t, *J* = 7.3 Hz, 3H);

^{13}C-NMR (100 MHz, $CDCl_3$) δ 188.2, 156.5, 138.7, 129.1, 125.5, 124.6, 96.7, 32.4, 27.8, 22.0, 19.7, 13.6;

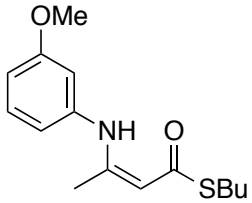
HRMS (ESI) m/z calcd for $C_{14}H_{19}NOS$ $(M+H)^+$ 250.1266; found 250.1277.





Compound **21d**

S-Butyl (*Z*)-3-((3-methoxyphenyl)amino)but-2-enethioate

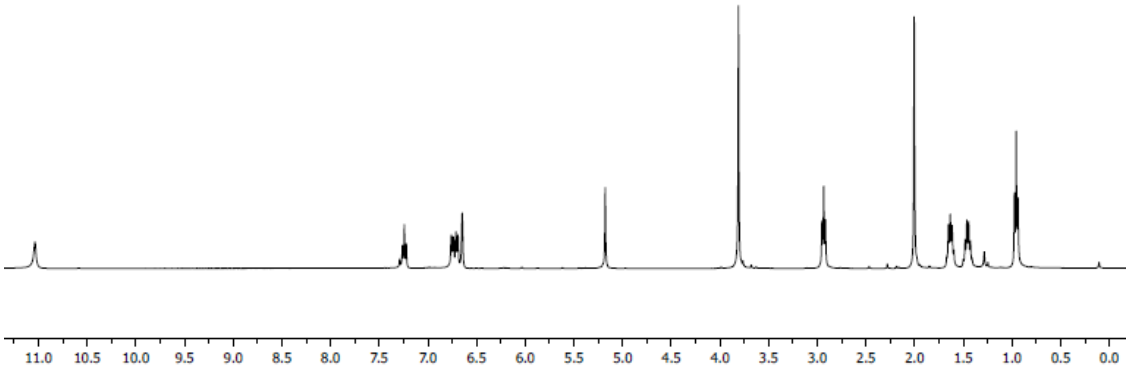


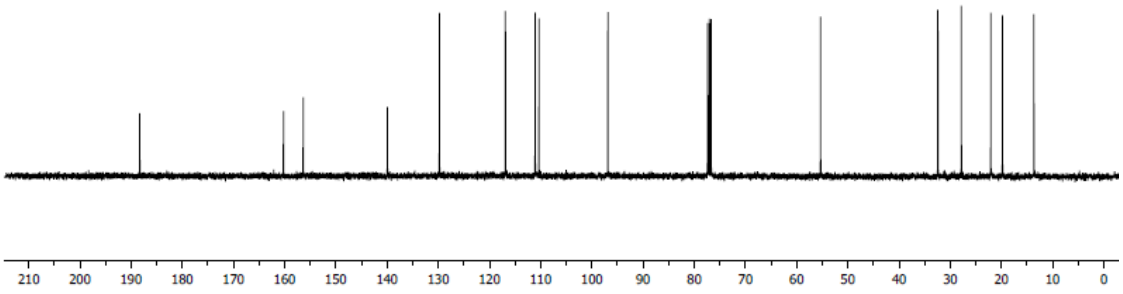
Colorless oil, 88 %;

^{1}H-NMR (400 MHz, $CDCl_3$) δ 11.04 (s, 1H), 7.24 (t, J = 8.1 Hz, 1H), 6.74 (d, J = 8.0 Hz, 1H), 6.70 (d, J = 8.0 Hz, 1H), 6.65 (s, 1H), 5.18 (s, 1H), 3.81 (s, 3H), 2.93 (t, J = 7.3 Hz, 2H), 2.00 (s, 3H), 1.69-1.55 (m, 2H), 1.49-1.40 (m, 2H), 0.96 (t, J = 7.3 Hz, 3H);

^{13}C-NMR (100 MHz, $CDCl_3$) δ 188.3, 160.2, 156.3, 139.9, 129.7, 116.8, 111.0, 110.3, 96.8, 55.3, 32.4, 27.8, 22.0, 19.8, 13.6;

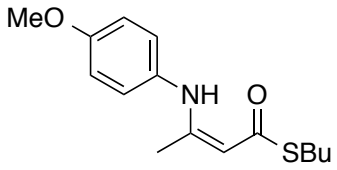
HRMS (ESI) m/z calcd for $C_{15}H_{21}NO_2S$ $(M+H)^+$ 280.1371; found 280.1385.





Compound **21e**

S-butyl (*Z*)-3-((4-methoxyphenyl)amino)but-2-enethioate

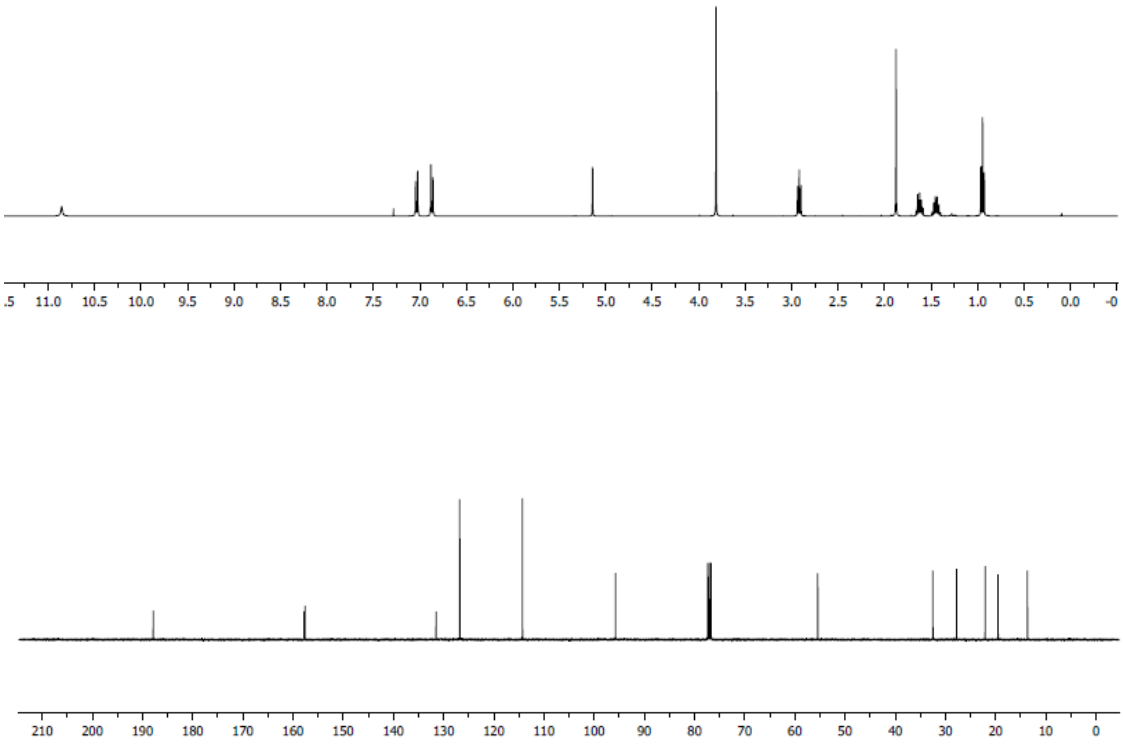


Colorless oil, 93 %;

^{1}H-NMR (400 MHz, $CDCl_3$) δ 10.85 (s, 1H), 7.04 (d, *J* = 8.8 Hz, 2H), 6.87 (d, *J* = 8.9 Hz, 2H), 5.14 (s, 1H), 3.81 (s, 3H), 2.92 (t, *J* = 7.3 Hz, 2H), 1.88 (s, 3H), 1.69-1.56 (m, 2H), 1.49-1.40 (m, 2H), 1.28 (s, 1H), 0.95 (t, *J* = 7.3 Hz, 3H);

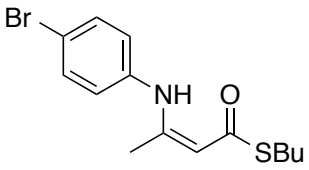
^{13}C-NMR (100 MHz, $CDCl_3$) δ 187.8, 157.7, 157.6, 131.5, 126.7, 114.2, 95.7, 55.4, 32.4, 27.7, 22.0, 19.5, 13.6;

HRMS (ESI+) m/z calcd for $C_{15}H_{21}NO_2S$ $(M+H)^+$ 280.1371; found 280.1379.



Compound **21f**

S-Butyl (*Z*)-3-((4-bromophenyl)amino)but-2-enethioate

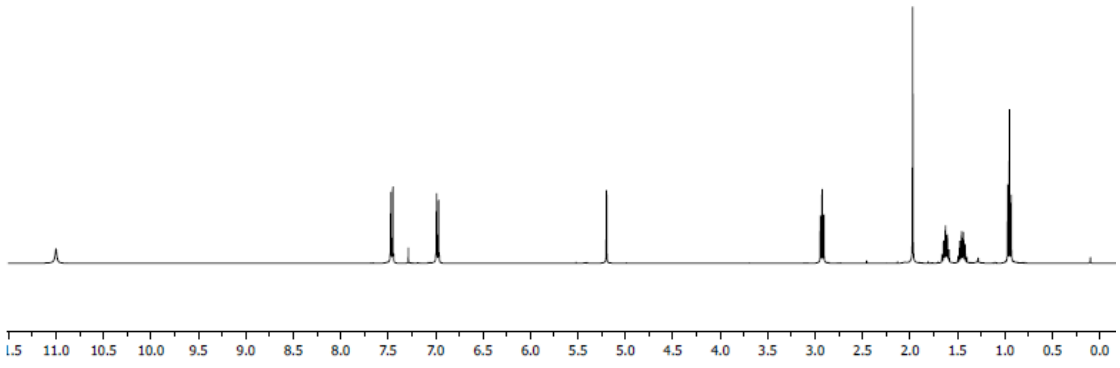


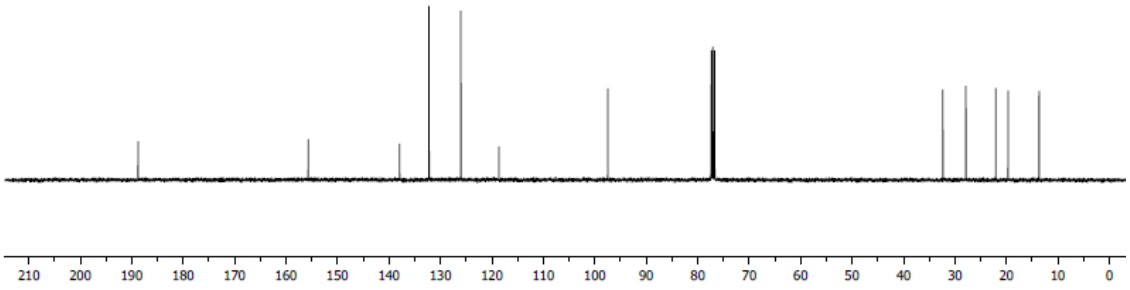
White solid, 91 %; mp = 55.0-56.1 °C;

^{1}H-NMR (400 MHz, $CDCl_3$) δ 11.00 (s, 1H), 7.50-7.43 (m, 2H), 7.01-6.94 (m, 2H), 5.20 (s, 1H), 2.93 (t, *J* = 7.3 Hz, 2H), 1.97 (s, 3H), 1.66-1.59 (m, 2H), 1.49-1.40 (m, 2H), 0.95 (t, *J* = 7.3 Hz, 3H);

^{13}C-NMR (100 MHz, $CDCl_3$) δ 188.7, 155.6, 137.9, 132.2, 126.0, 118.6, 97.4, 32.3, 27.8, 22.0, 19.6, 13.6;

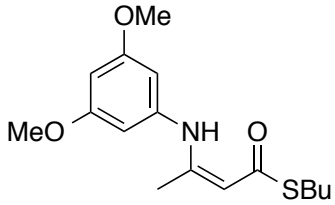
HRMS (ESI) m/z calcd for $C_{14}H_{18}BrNOS$ $(M+H)^+$ 328.0371; found 328.0386.





Compound **21g**

S-Butyl (*Z*)-3-((3,5-dimethoxyphenyl)amino)but-2-enethioate

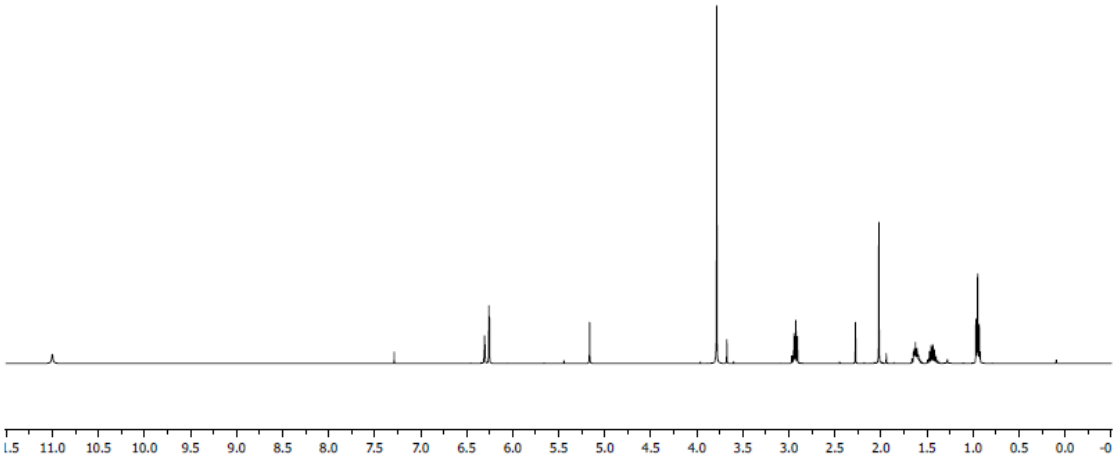


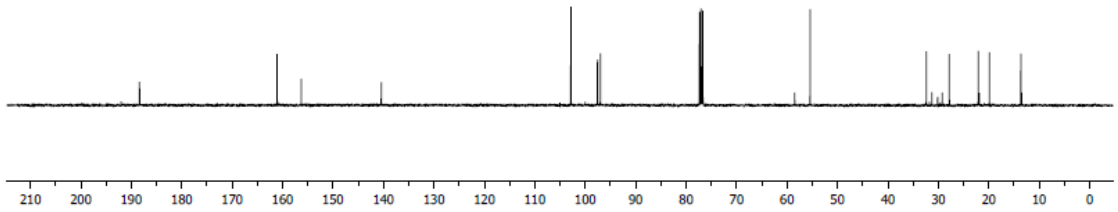
Light yellow oil, 89 %;

^{1}H-NMR (400 MHz, $CDCl_3$) δ 11.00 (s, 1H), 6.31 (t, *J* = 2.2 Hz, 1H), 6.26 (d, *J* = 2.1 Hz, 2H), 5.16 (s, 1H), 3.78 (s, 6H), 2.93 (t, *J* = 7.4 Hz, 2H), 2.02 (s, 3H), 1.69-1.54 (m, 2H), 1.52-1.34 (m, 2H), 0.95 (t, *J* = 7.3 Hz, 3H);

^{13}C-NMR (100 MHz, $CDCl_3$) δ 188.3, 161.0, 156.2, 140.4, 102.8, 97.5, 96.9, 55.4, 32.3, 27.8, 22.0, 19.8, 13.6;

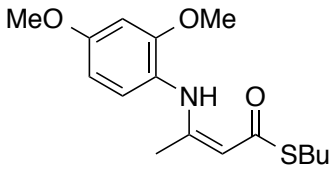
HRMS (ESI+) m/z calcd for $C_{16}H_{23}NO_3S$ $(M+H)^+$ 310.1477; found 310.1486.





Compound **21h**

S-Butyl (*Z*)-3-((2,4-dimethoxyphenyl)amino)but-2-enethioate

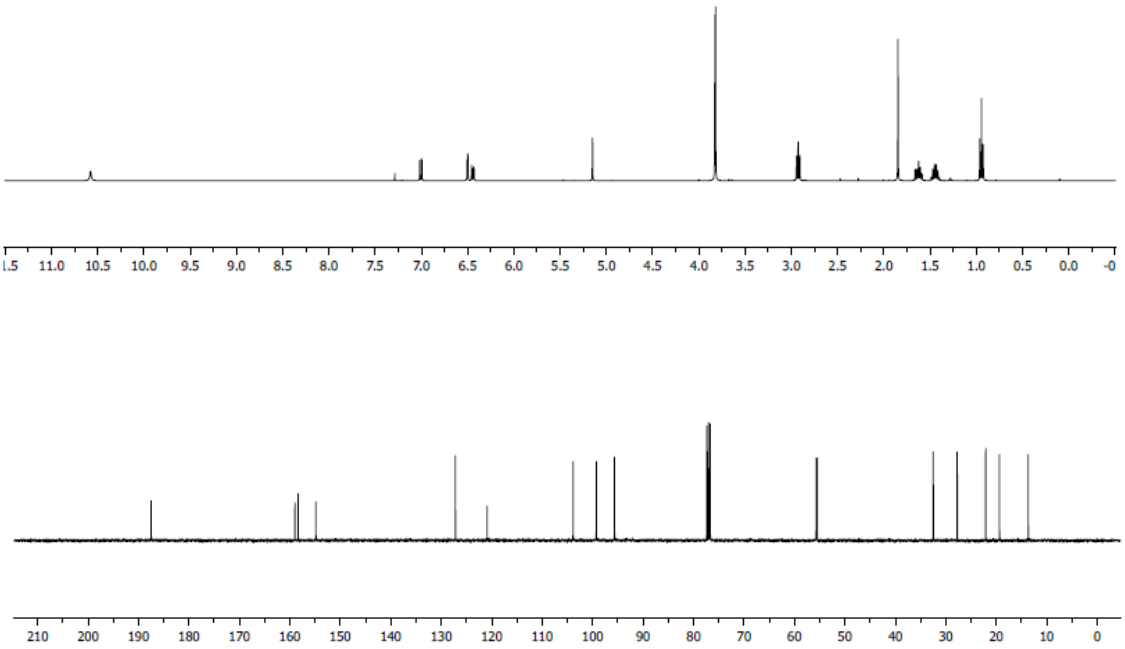


Colorless oil, 92 %;

^{1}H-NMR (400 MHz, $CDCl_3$) δ 10.58 (s, 1H), 7.01 (d, *J* = 8.6 Hz, 1H), 6.50 (d, *J* = 2.6 Hz, 1H), 6.44 (dd, *J* = 8.6, 2.6 Hz, 1H), 5.15 (s, 1H), 3.82 (d, *J* = 2.7 Hz, 6H), 2.93 (t, *J* = 7.3 Hz, 2H), 1.85 (s, 3H), 1.67-1.58 (m, 2H), 1.49-1.41 (m, 2H), 0.94 (t, *J* = 7.3 Hz, 3H);

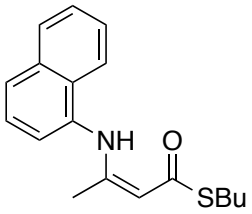
^{13}C-NMR (100 MHz, $CDCl_3$) δ 187.4, 159.0, 158.3, 154.8, 127.1, 120.9, 103.8, 99.2, 95.6, 55.6, 55.4, 32.4, 27.7, 22.1, 19.3, 13.6;

HRMS (ESI+) m/z calcd for $C_{16}H_{23}NO_3S$ $(M+H)^+$ 310.1477; found 310.1467.



Compound **21i**

S-Butyl (*Z*)-3-(naphthalen-1-ylamino)but-2-enethioate

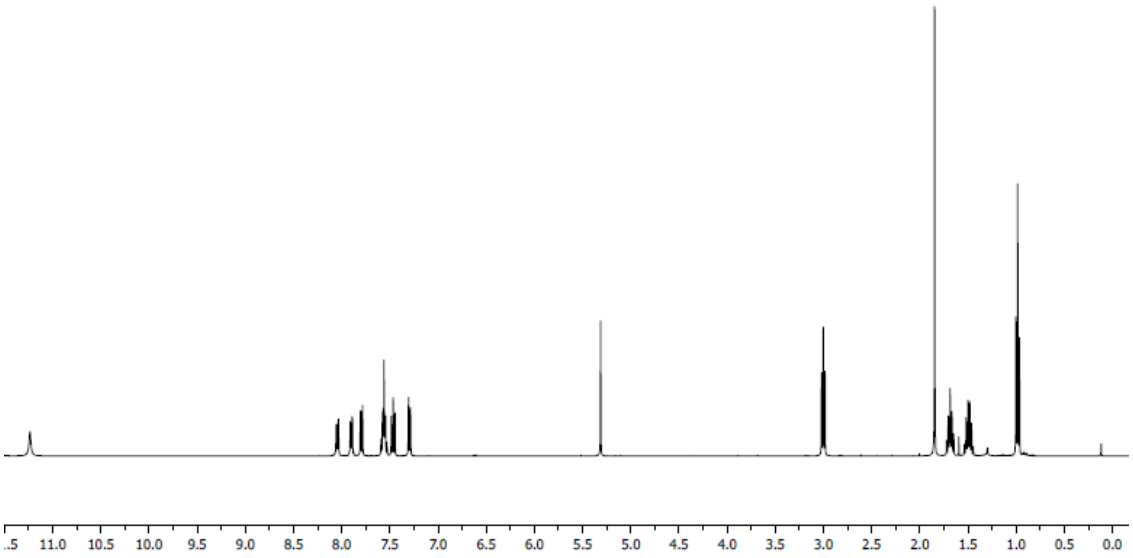


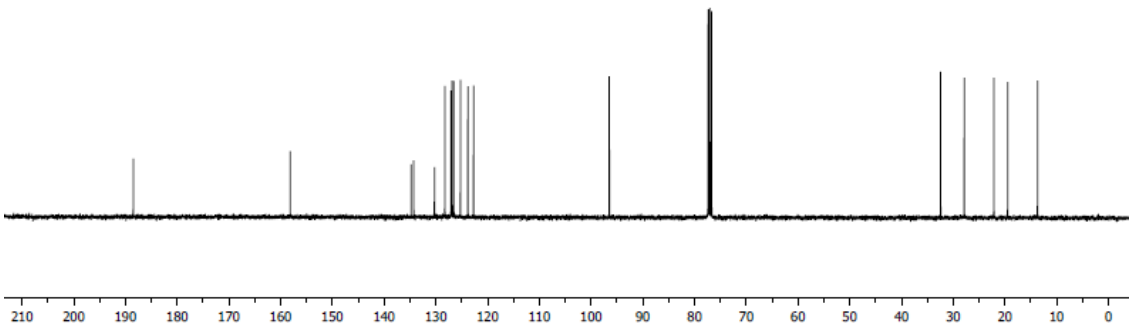
Colorless oil, 85 %;

^{1}H-NMR (400 MHz, $CDCl_3$) δ 11.24 (s, 1H), 8.04 (dd, *J* = 8.1, 1.4 Hz, 1H), 7.90 (dd, *J* = 6.6, 2.8 Hz, 1H), 7.79 (d, *J* = 8.2 Hz, 1H), 7.60-7.52 (m, 2H), 7.51-7.42 (m, 1H), 7.33-7.26 (m, 1H), 5.30 (s, 1H), 3.00 (t, *J* = 7.3 Hz, 2H), 1.85 (s, 3H), 1.72-1.65 (m, 2H), 1.54-1.46 (m, 2H), 0.98 (t, *J* = 7.3 Hz, 3H);

^{13}C-NMR (100 MHz, $CDCl_3$) δ 188.4, 158.1, 134.7, 134.2, 130.2, 128.2, 127.0, 126.9, 126.5, 125.2, 123.8, 122.7, 96.4, 32.4, 27.8, 22.1, 19.4, 13.6;

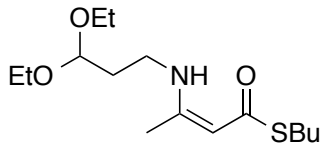
HRMS (ESI) m/z calcd for $C_{18}H_{21}NOS$ $(M+H)^+$ 300.1422; found 300.1437.





Compound **21j**

S-Butyl (*Z*)-3-((3,3-diethoxypropyl)amino)but-2-enethioate

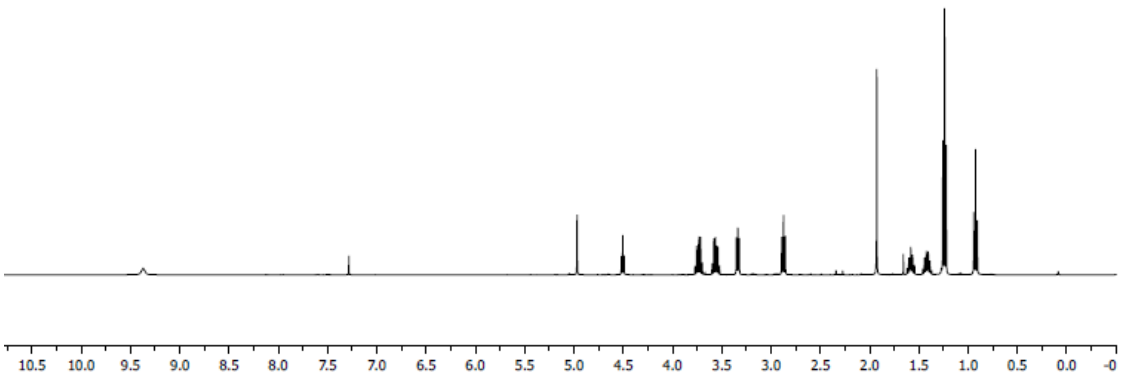


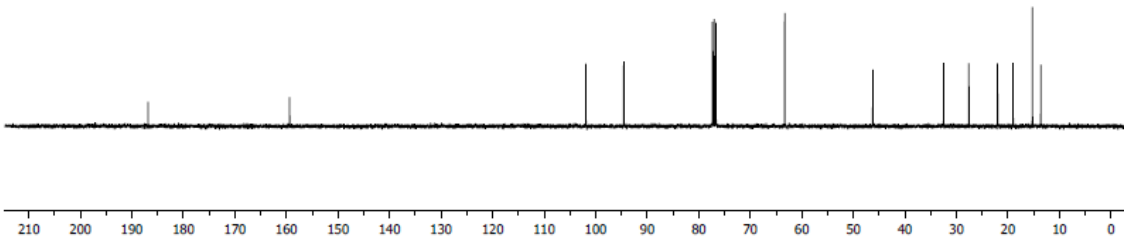
Colorless oil, 95 %;

^{1}H-NMR (400 MHz, $CDCl_3$) δ 9.37 (s, 1H), 4.97 (s, 1H), 4.50 (t, *J* = 5.5 Hz, 1H), 3.79-3.68 (m, 2H), 3.60-3.49 (m, 2H), 3.34 (t, *J* = 6.0 Hz, 2H), 2.87 (t, *J* = 7.3 Hz, 2H), 1.93 (s, 3H), 1.68-1.52 (m, 2H), 1.46-1.38 (m, 2H), 1.24 (t, *J* = 7.4 Hz, 6H), 0.93 (t, *J* = 7.3 Hz, 3H);

^{13}C-NMR (100 MHz, $CDCl_3$) δ 186.8, 159.3, 101.9, 94.5, 63.3, 46.2, 32.5, 27.5, 22.0, 19.0, 15.2, 13.6;

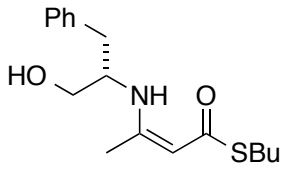
HRMS (ESI+) m/z calcd for $C_{15}H_{29}NO_3S$ $(M+H)^+$ 304.1946; found 304.1955.





Compound **21k**

S-Butyl (*S*,*Z*)-3-((1-hydroxy-3-phenylpropan-2-yl)amino)but-2-enethioate

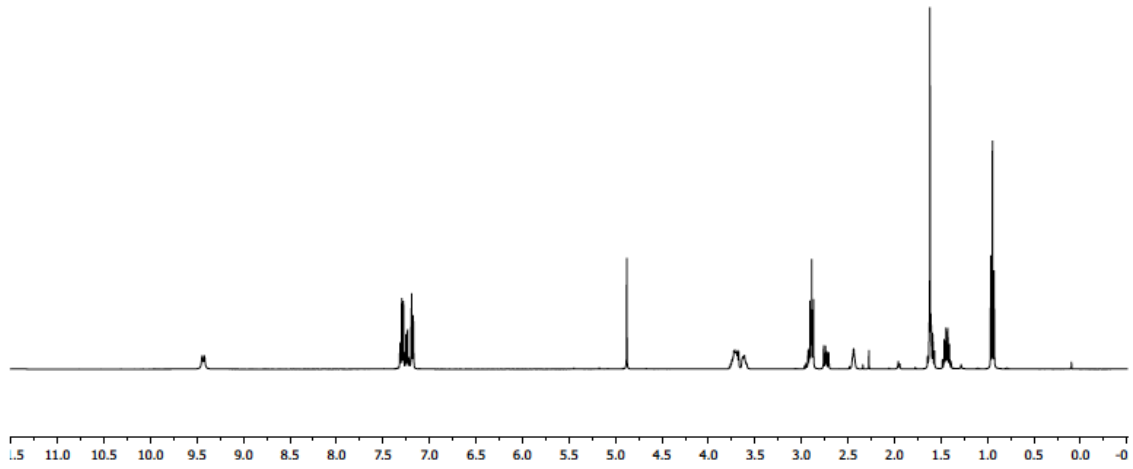


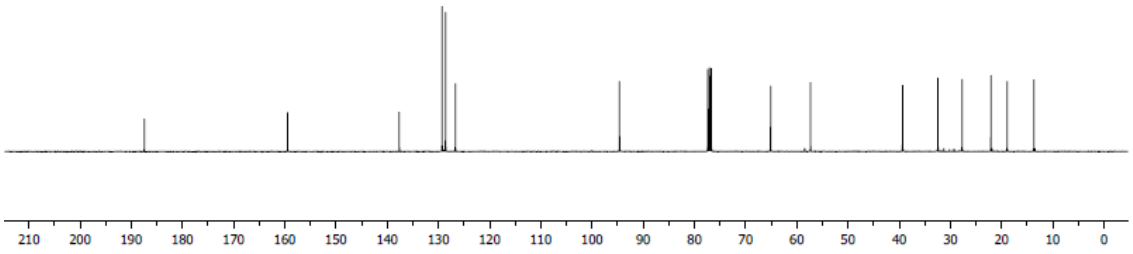
Colorless oil, 94 %;

^{1}H-NMR (400 MHz, $CDCl_3$) δ 9.43 (d, *J* = 9.8 Hz, 1H), 7.35-7.15 (m, 5H), 4.88 (s, 1H), 3.90-3.52 (m, 3H), 2.96-2.85 (m, 3H), 2.73 (dd, *J* = 13.6, 8.1 Hz, 1H), 2.44 (s, 1H), 1.64-1.54 (m, 5H), 1.51-1.36 (m, 2H), 0.95 (t, *J* = 7.4 Hz, 3H);

^{13}C-NMR (100 MHz, $CDCl_3$) δ 187.4, 159.4, 137.6, 129.2, 128.6, 126.6, 94.6, 65.1, 57.3, 39.3, 32.4, 27.7, 22.0, 18.9, 13.6;

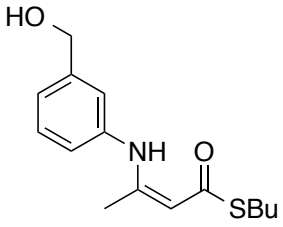
HRMS (ESI) m/z calcd for $C_{17}H_{25}NO_2S$ $(M+H)^+$ 308.1684; found 308.1670.





Compound **21l**

S-butyl (*Z*)-3-((3-(hydroxymethyl)phenyl)amino)but-2-enethioate

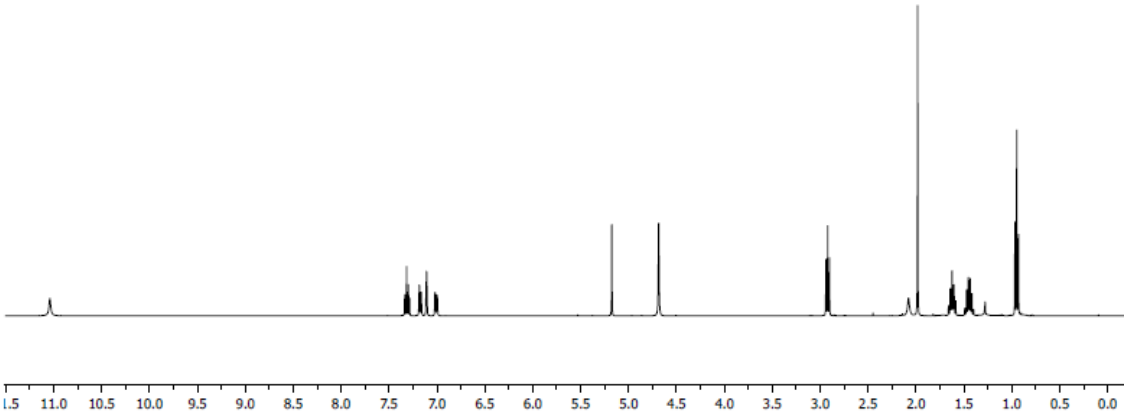


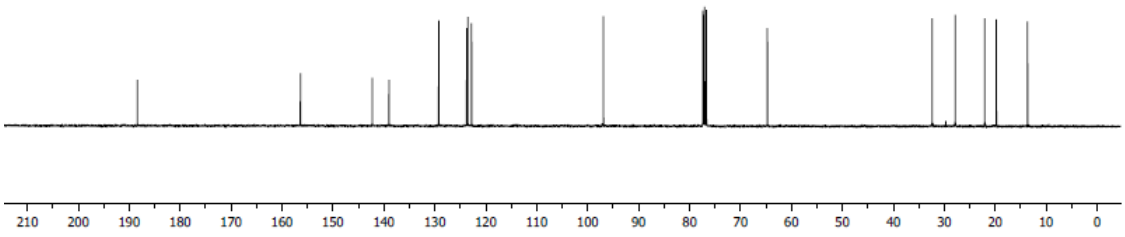
Colorless oil, 78 % (this compound was purified by flash chromatography CH_2Cl_2 to 5 % EtOAc in CH_2Cl_2);

^{1}H-NMR (400 MHz, $CDCl_3$) δ 11.04 (s, 1H), 7.32 (t, *J* = 7.8 Hz, 1H), 7.17 (d, *J* = 7.7 Hz, 1H), 7.11 (s, 1H), 7.01 (d, *J* = 7.9 Hz, 1H), 5.18 (s, 1H), 4.69 (s, 2H), 2.92 (t, *J* = 7.3 Hz, 2H), 2.08 (s, 1H), 1.98 (s, 3H), 1.64-1.58 (m, 2H), 1.47-1.42 (m, 2H), 0.95 (t, *J* = 7.3 Hz, 3H);

^{13}C-NMR (100 MHz, $CDCl_3$) δ 188.3, 156.3, 142.2, 139.0, 129.2, 123.7, 123.5, 122.7, 96.9, 64.7, 32.3, 27.8, 22.0, 19.7, 13.6;

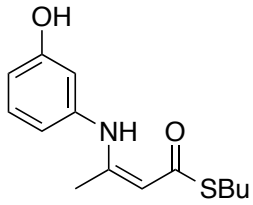
HRMS (ESI) m/z calcd for $C_{15}H_{21}NO_2S$ $(M+Li)^+$ 286.1453; found 286.1440.





Compound **21m**

S-Butyl (*Z*)-3-((3-hydroxyphenyl)amino)but-2-enethioate

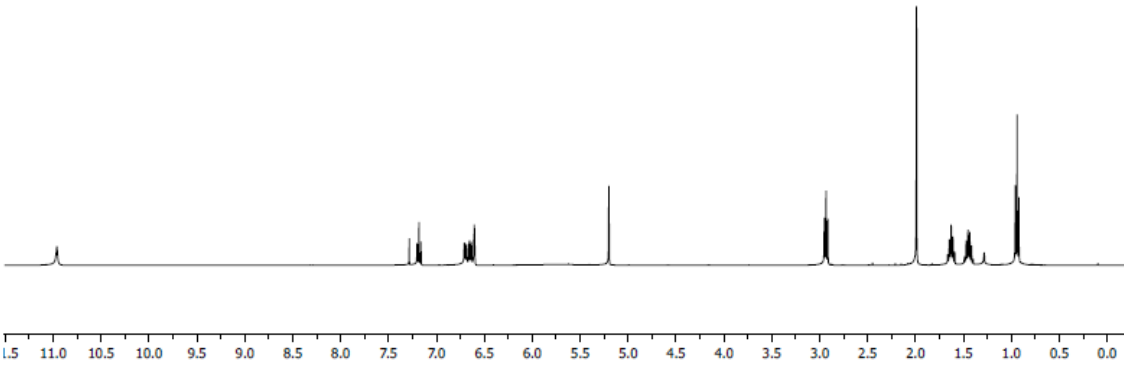


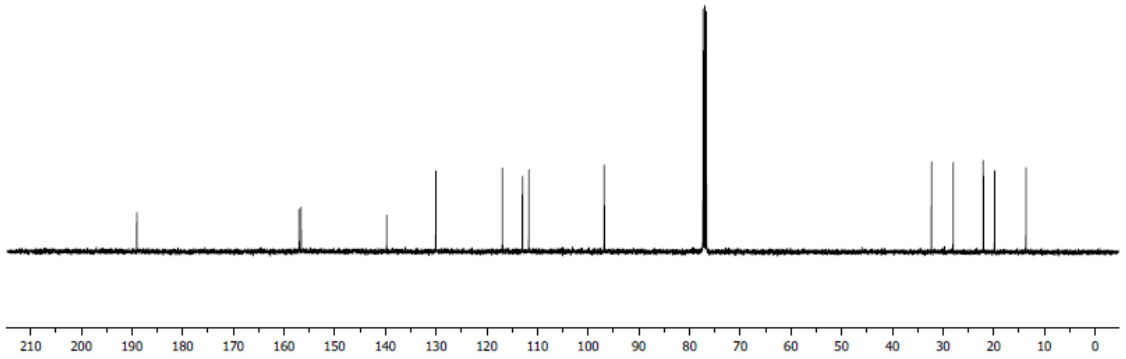
Light yellow oil, 51 % (this compound was purified by flash chromatography, CH_2Cl_2 to 5 % EtOAc in CH_2Cl_2);

^{1}H-NMR (400 MHz, $CDCl_3$) δ 10.96 (s, 1H), 7.18 (t, *J* = 8.0 Hz, 1H), 6.79-6.57 (m, 3H), 5.20 (s, 1H), 2.93 (t, *J* = 7.3 Hz, 2H), 1.99 (s, 3H), 1.65-1.61 (m, 2H), 1.49-1.42 (m, 2H), 0.94 (t, *J* = 7.3 Hz, 3H);

^{13}C-NMR (100 MHz, $CDCl_3$) δ 189.0, 157.0, 156.6, 139.7, 130.0, 116.9, 112.9, 111.7, 96.8, 32.2, 27.9, 22.0, 19.8, 13.6;

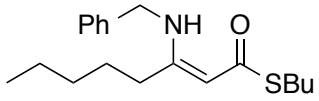
HRMS (ESI) m/z calcd for $C_{14}H_{19}NO_2S$ $(M+H)^+$ 266.1215; found 266.1226.





Compound **21n**

S-Butyl (*Z*)-3-(benzylamino)oct-2-enethioate

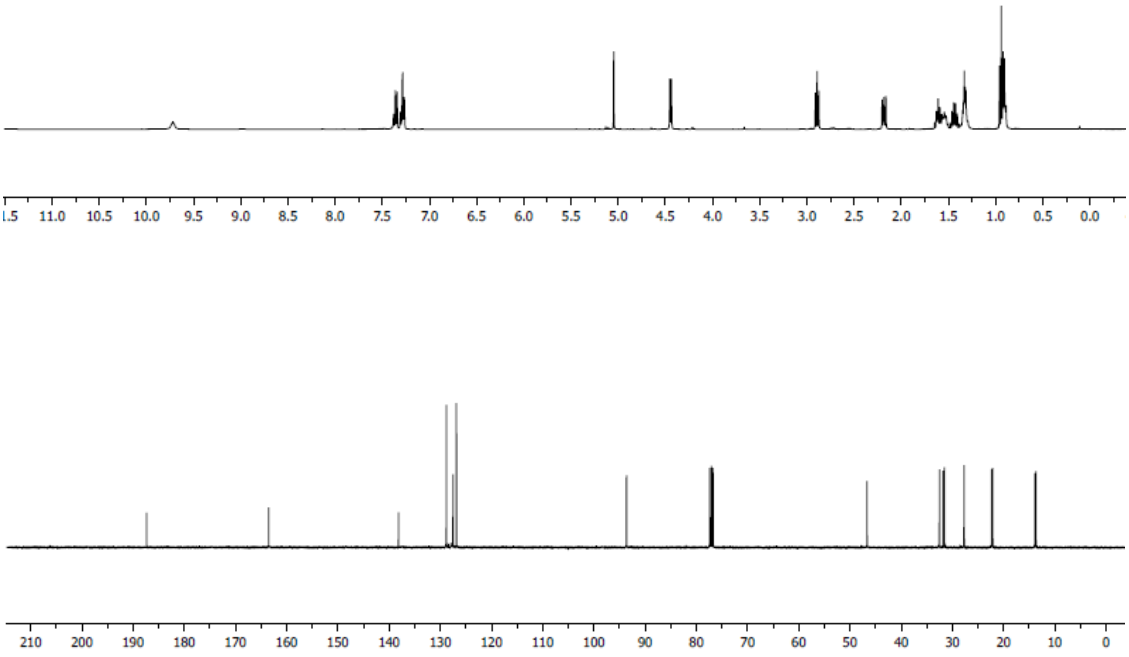


Light yellow oil, 85 %;

^{1}H-NMR (400 MHz, $CDCl_3$) δ 9.72 (s, 1H), 7.42-7.27 (m, 5H), 5.05 (s, 1H), 4.44 (d, *J* = 6.4 Hz, 2H), 2.89 (t, *J* = 7.3 Hz, 2H), 2.23-2.14 (t, *J* = 7.8 Hz, 2H), 1.66-1.49 (m, 4H), 1.46-1.40 (m, 2H), 1.38-1.14 (m, 4H), 0.98-0.83 (m, 6H);

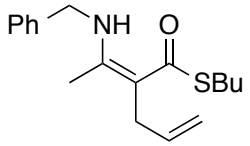
^{13}C-NMR (100 MHz, $CDCl_3$) δ 187.3, 163.5, 138.1, 128.8, 127.5, 126.8, 93.6, 46.6, 32.5, 31.7, 31.5, 27.7, 27.6, 22.3, 22.1, 13.8, 13.6;

HRMS (ESI) m/z calcd for $C_{19}H_{29}NOS$ $(M+H)^+$ 320.2048; found 320.2033.



Compound **21o**

S-Butyl (*Z*)-2-(1-(benzylamino)ethylidene)pent-4-enethioate

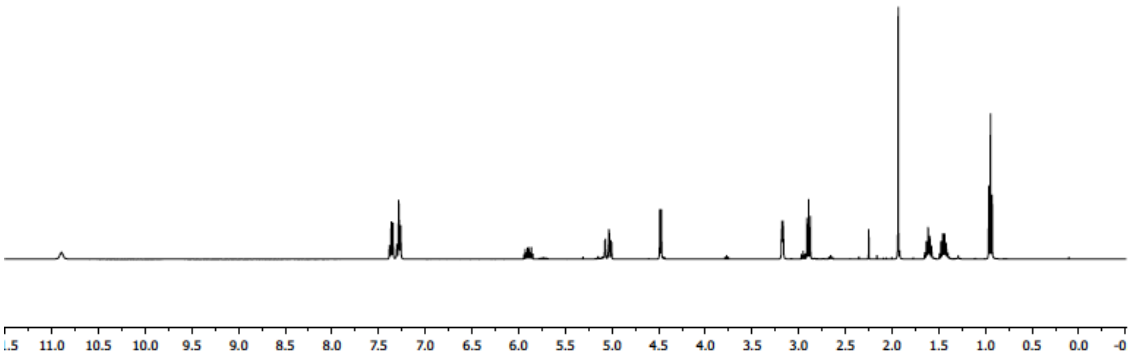


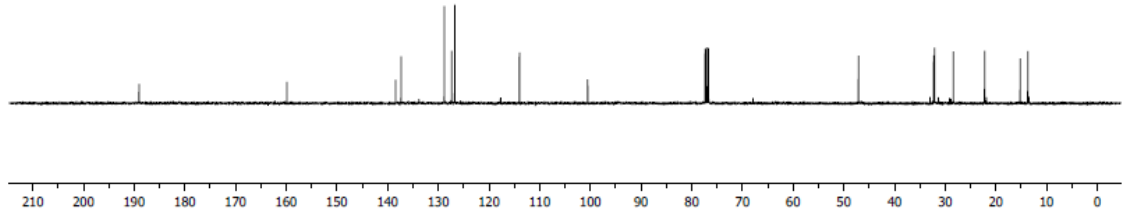
Light yellow oil, 74 %;

^{1}H-NMR (400 MHz, $CDCl_3$) δ 10.90 (s, 1H), 7.42-7.23 (m, 5H), 5.95-5.82 (m, 1H), 5.25-4.97 (m, 2H), 4.48 (d, J = 6.2 Hz, 2H), 3.17 (dt, J = 5.4, 1.7 Hz, 2H), 2.89 (t, J = 7.4 Hz, 2H), 1.93 (s, 3H), 1.65-1.56 (m, 2H), 1.49-1.40 (m, 2H), 0.95 (t, J = 7.3 Hz, 3H);

^{13}C-NMR (100 MHz, $CDCl_3$) δ 189.0, 159.8, 138.4, 137.3, 128.8, 127.3, 126.7, 113.9, 100.5, 47.1, 32.2, 32.1, 28.3, 22.2, 15.2, 13.7;

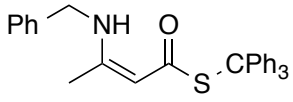
HRMS (ESI) m/z calcd for $C_{18}H_{25}NOS$ $(M+H)^+$ 304.1735; found 304.1747.





Compound **21p**

S-Trityl (*Z*)-3-(benzylamino)but-2-enethioate

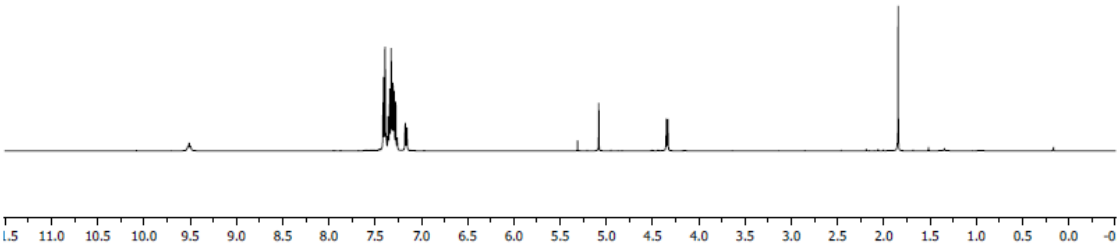


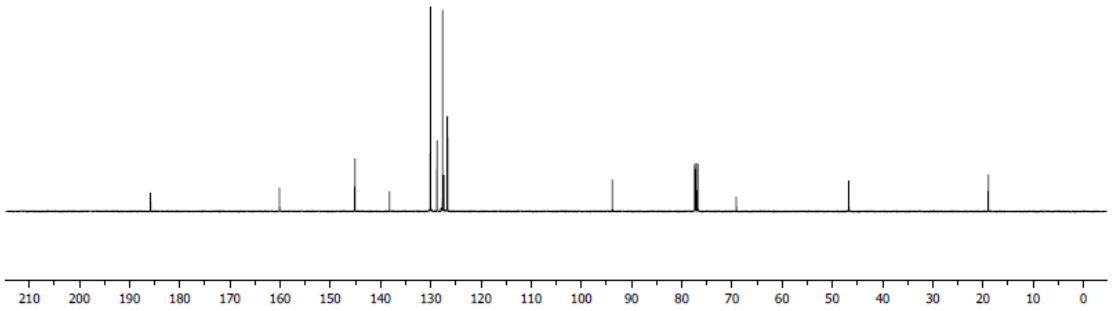
Light yellow solid, 91 %; mp = 151.2-152.2 °C;

^{1}H-NMR (400 MHz, $CDCl_3$) δ 9.51 (t, *J* = 6.7 Hz, 1H), 7.44-7.23 (m, 18H), 7.21-7.13 (m, 2H), 5.08 (s, 1H), 4.34 (d, *J* = 6.7 Hz, 2H), 1.85 (s, 3H);

^{13}C-NMR (100 MHz, $CDCl_3$) δ 185.8, 160.1, 145.1, 138.2, 130.0, 128.7, 127.6, 127.3, 126.7, 126.6, 93.8, 69.1, 46.7, 18.9;

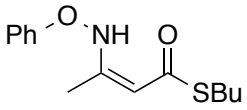
HRMS (MALDI+) m/z calcd for $C_{30}H_{27}NOS$ $(M+H)^+$ 450.1892; found 450.1879.





Compound **21q**

S-Butyl (*Z*)-3-(phenoxyamino)but-2-enethioate

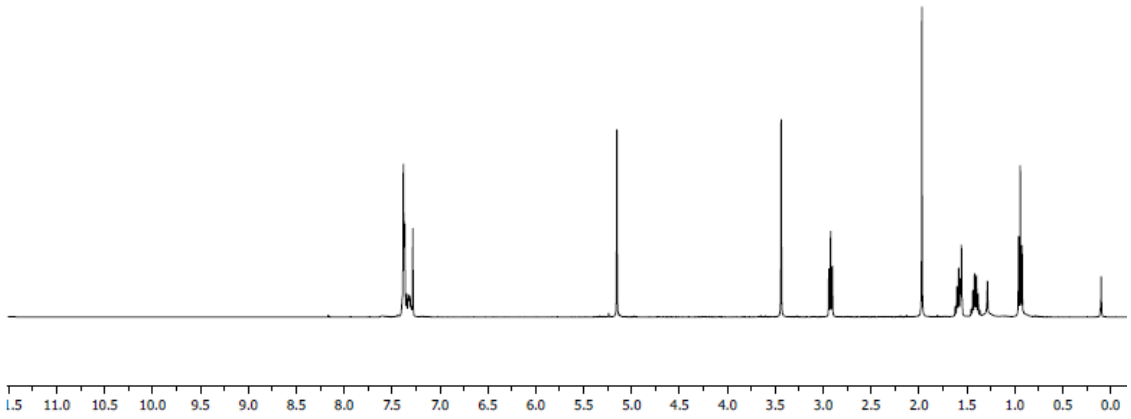


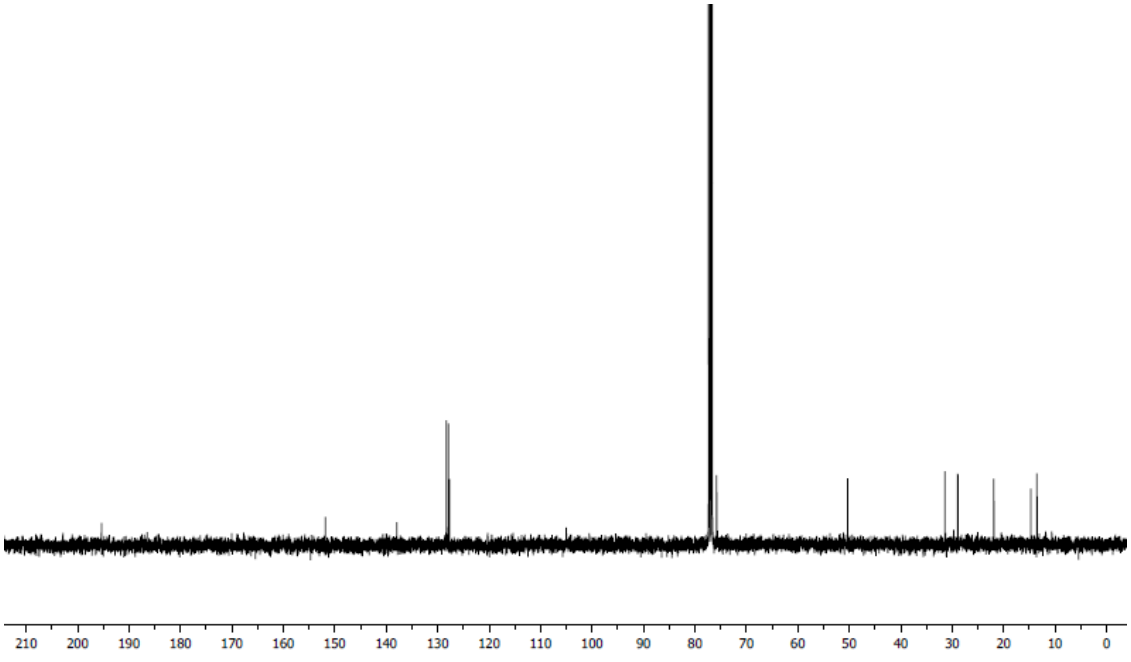
Colorless oil, 58 %;

^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.42-7.26 (m, 5H), 5.15 (s, 2H), 3.44 (s, 2H), 2.92 (t, *J* = 7.3 Hz, 2H), 1.97 (s, 3H), 1.61-1.50 (m, 2H), 1.46-1.38 (m, 2H), 0.94 (t, *J* = 7.3 Hz, 3H);

^{13}C-NMR (100 MHz, $CDCl_3$) δ 195.3, 151.8, 137.9, 128.3, 127.9, 127.6, 75.7, 50.3, 31.4, 28.9, 21.9, 14.7, 13.5;

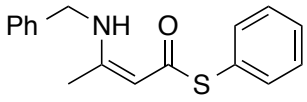
HRMS (ESI) m/z calcd for $C_{15}H_{21}NO_2S$ $(M+H)^+$ 280.1371; found 280.1381.





Compound **21r**

S-Phenyl (*Z*)-3-(benzylamino)but-2-enethioate

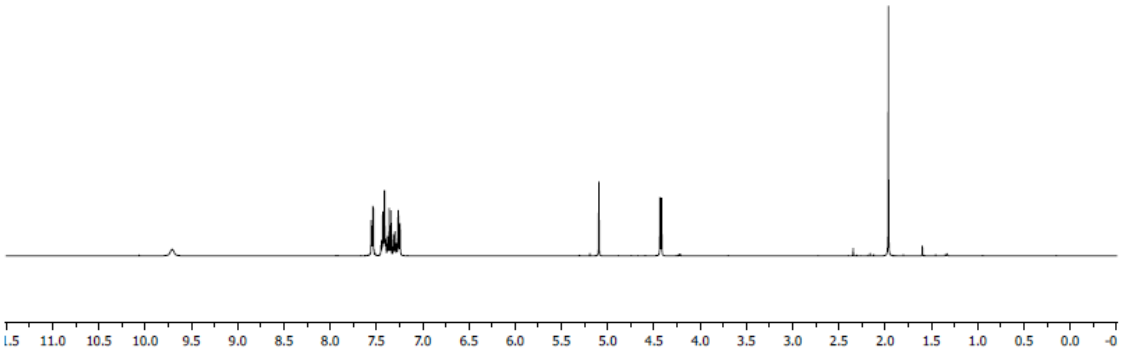


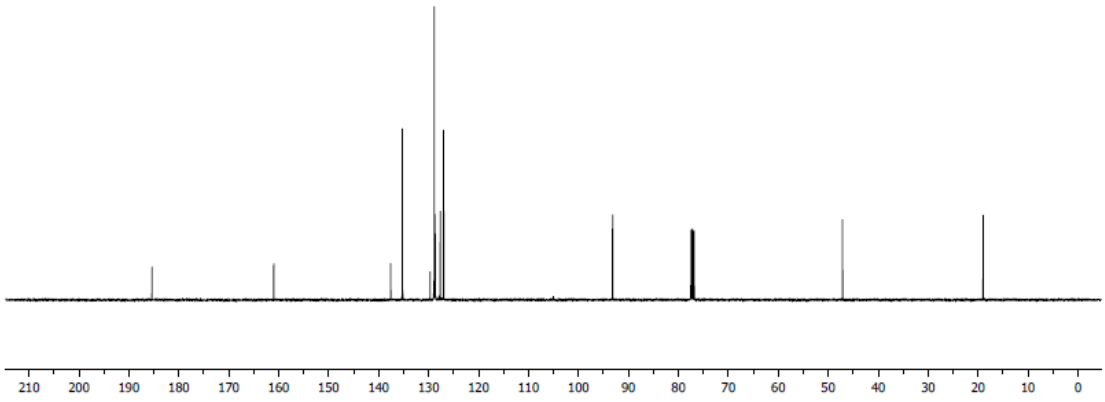
Colorless oil, 76 %;

^{1}H-NMR (400 MHz, $CDCl_3$) δ 9.71 (s, 1H), 7.59-7.50 (m, 2H), 7.47-7.22 (m, 8H), 5.09 (s, 1H), 4.43 (d, J = 6.3 Hz, 2H), 1.97 (s, 3H);

^{13}C-NMR (100 MHz, $CDCl_3$) δ 185.3, 161.0, 137.5, 135.2, 129.7, 128.8, 128.7, 127.6, 126.9, 93.1, 47.1, 18.9;

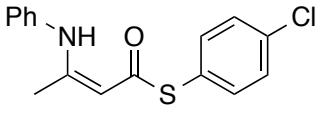
HRMS (ESI) m/z calcd for $C_{18}H_{25}NOS$ $(M+H)^+$ 284.1109; found 284.1101.





Compound **21t**

S-(4-Chlorophenyl) (*Z*)-3-(phenylamino)but-2-enethioate

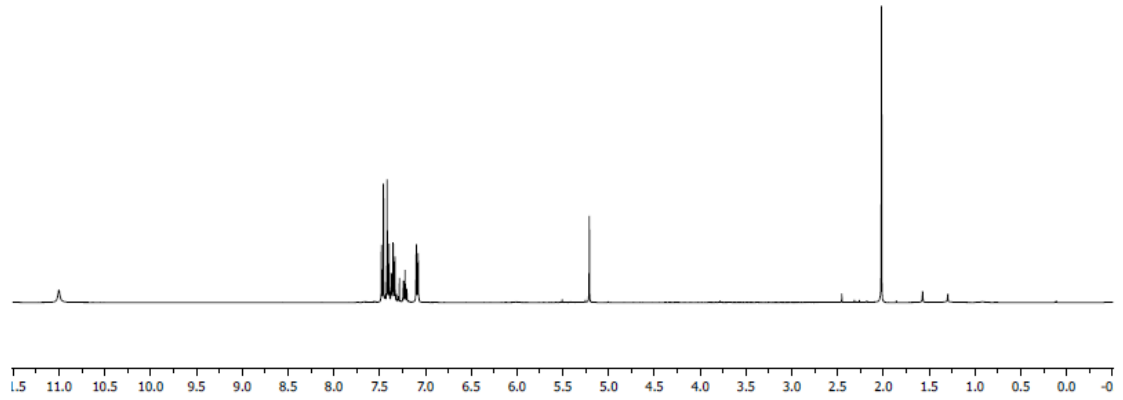


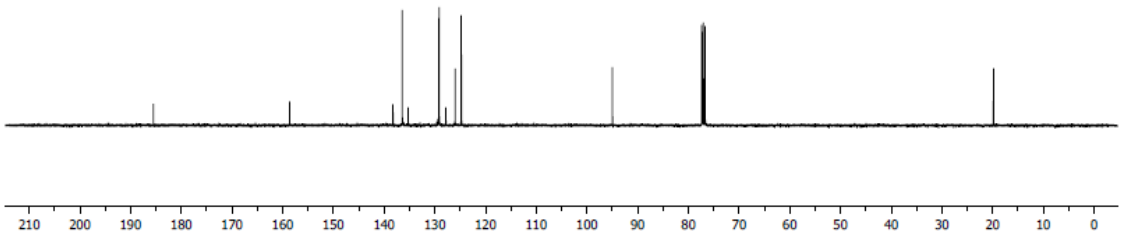
Light yellow solid, 72 %; mp = 75.1-76.9 °C;

^{1}H-NMR (400 MHz, $CDCl_3$) δ 11.00 (s, 1H), 7.51-7.18 (m, 8H), 7.09 (d, *J* = 7.5 Hz, 2H), 5.21 (s, 1H), 2.02 (s, 3H);

^{13}C-NMR (100 MHz, $CDCl_3$) δ 185.4, 158.6, 138.2, 136.4, 135.2, 129.2, 129.1, 127.8, 125.9, 124.7, 94.9, 19.8;

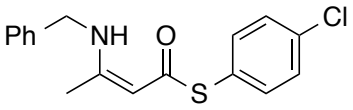
HRMS (ESI) m/z calcd for $C_{16}H_{14}ClNOS$ $(M+H)^+$ 304.0563; found 304.0560.





Compound **21s**

S-(4-Chlorophenyl) (*Z*)-3-(benzylamino)but-2-enethioate

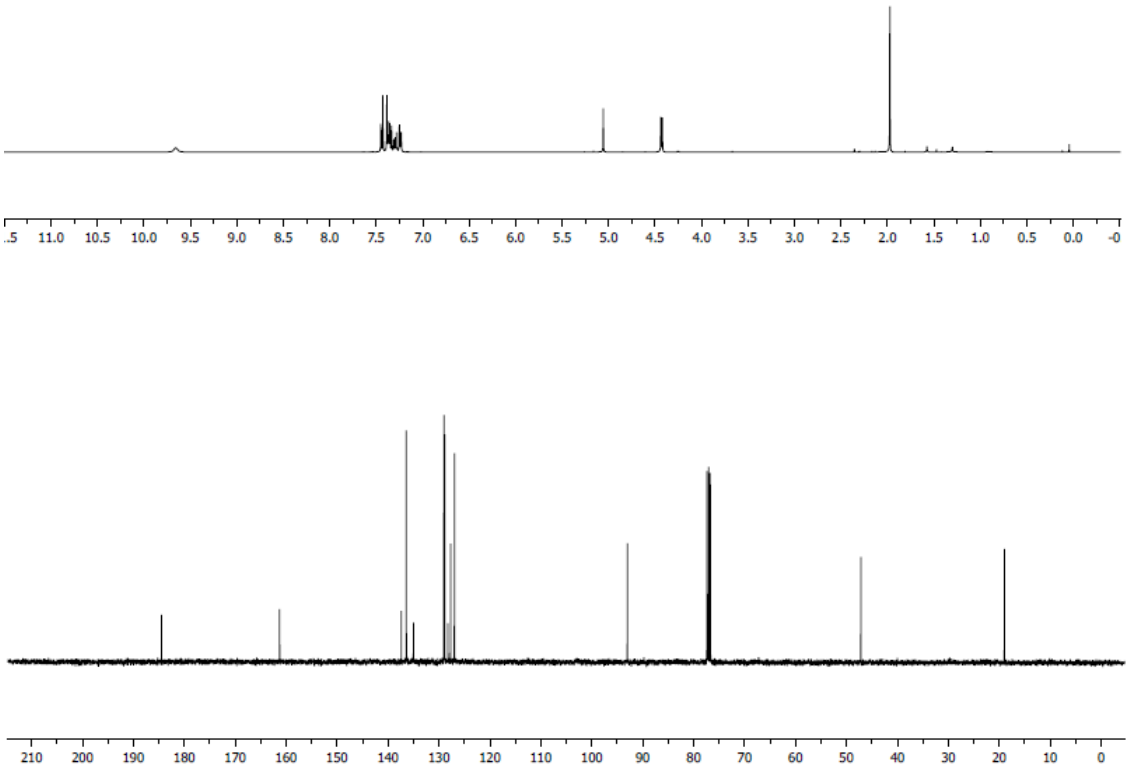


White solid, 80 %; mp = 94.5-96.0 °C;

^{1}H-NMR (400 MHz, $CDCl_3$) δ 9.66 (s, 1H), 7.48-7.20 (m, 9H), 5.06 (s, 1H), 4.43 (d, J = 6.3 Hz, 2H), 1.97 (s, 3H);

^{13}C-NMR (100 MHz, $CDCl_3$) δ 184.4, 161.2, 137.3, 136.4, 135.0, 129.0, 128.8, 128.2, 127.7, 126.9, 93.0, 47.1, 18.9;

HRMS (ESI+) m/z calcd for $C_{17}H_{16}ClNOS$ $(M+H)^+$ 318.0719; found 318.0724.

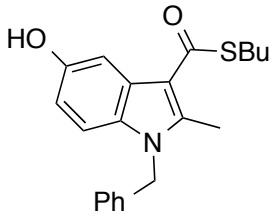


D. Syntheses of Indole Derivatives via Nenitzescu Reaction

A solution of **21** (0.3 mmol) and 1,4-benzoquinone (0.33 mmol) was added to 0.6 mL dry CH_3NO_2. The solution was stirred at 25 °C for 24 h. The solvent was removed under vacuum and the residue was purified by flash chromatography (10 % - 30 % EtOAc in hexanes) to give the desired product.

Compound **22a**

S-Butyl 1-benzyl-5-hydroxy-2-methyl-1*H*-indole-3-carbothioate



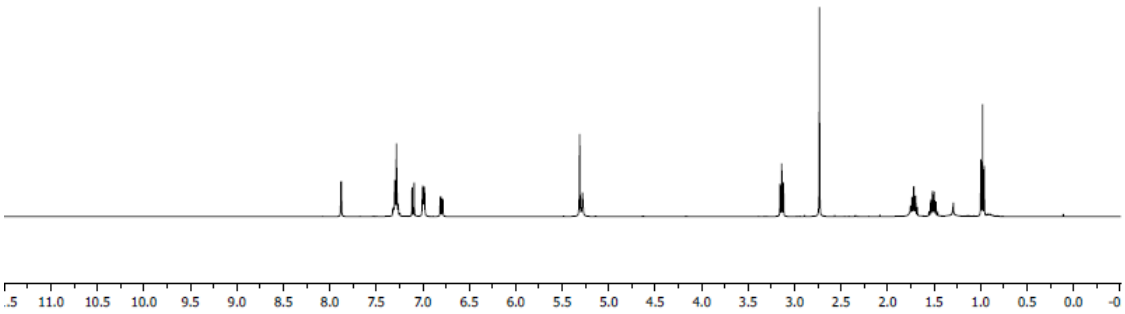
Light yellow solid, 75 %; mp = 150.6-151.3 °C;

^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.88 (d, *J* = 2.4 Hz, 1H), 7.35-7.26 (m, 3H), 7.10 (d, *J* = 8.7 Hz, 1H), 7.04-6.95 (m, 2H), 6.80 (dd, *J* = 8.7, 2.4 Hz, 1H), 5.31 (s, 2H), 5.28 (s, 1H), 3.14 (t, *J* = 7.3 Hz, 2H), 2.73 (s, 3H), 1.76-1.65 (m, 2H), 1.55-1.43 (m, 2H), 0.98 (t, *J* = 7.3 Hz, 3H);

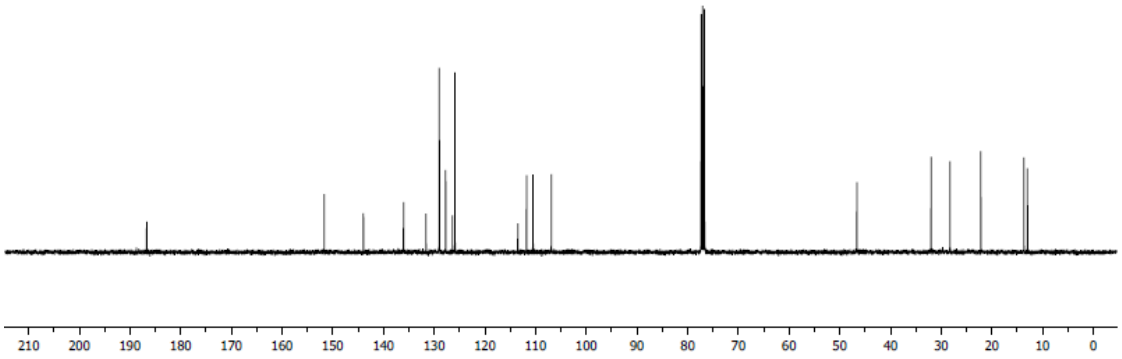
^{13}C-NMR (100 MHz, $CDCl_3$) δ 186.7, 151.6, 143.9, 136.0, 131.6, 128.9, 127.7, 126.3, 125.9, 113.5, 111.7, 110.5, 106.8, 46.6, 31.9, 28.2, 22.1, 13.6, 12.9;

HRMS (MALDI+) m/z calcd for $C_{21}H_{23}NO_2S$ $(M+H)^+$ 354.1522; found 354.1538.

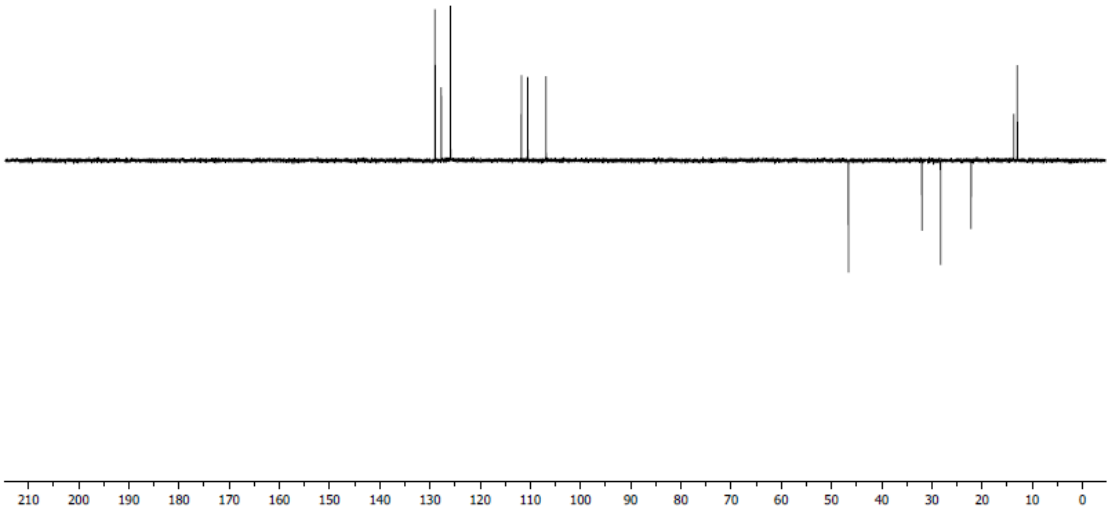
^{1}H NMR



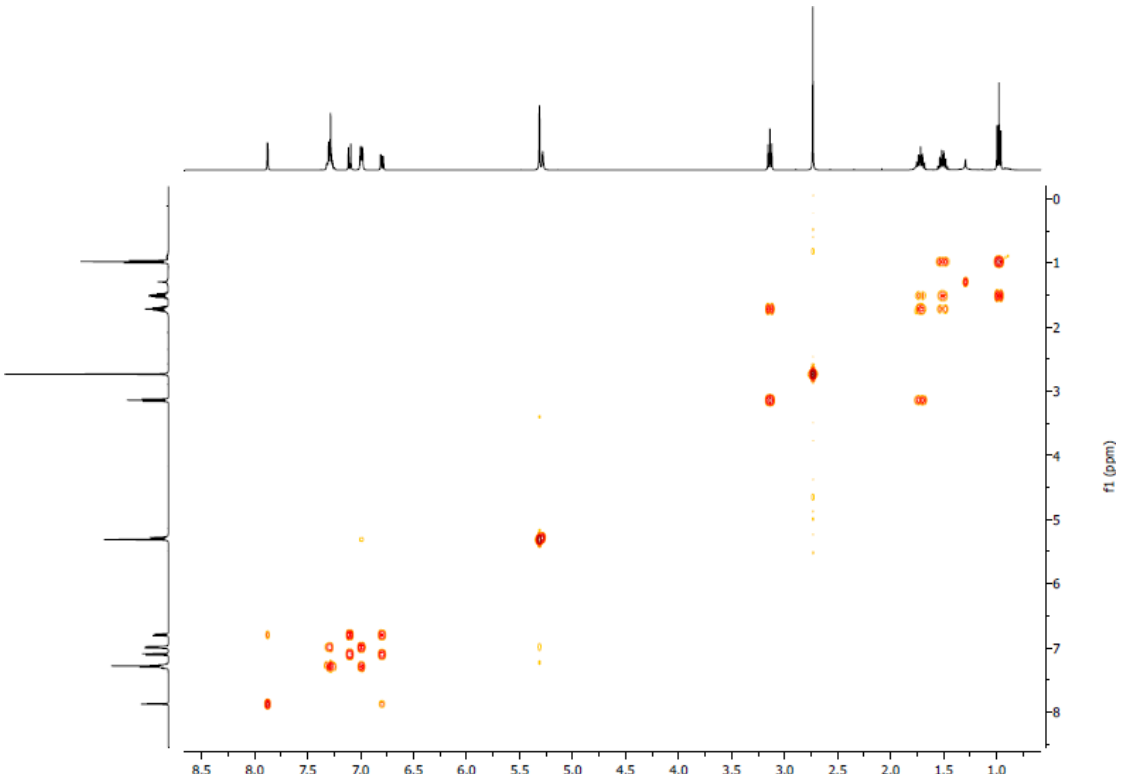
^{13}C NMR



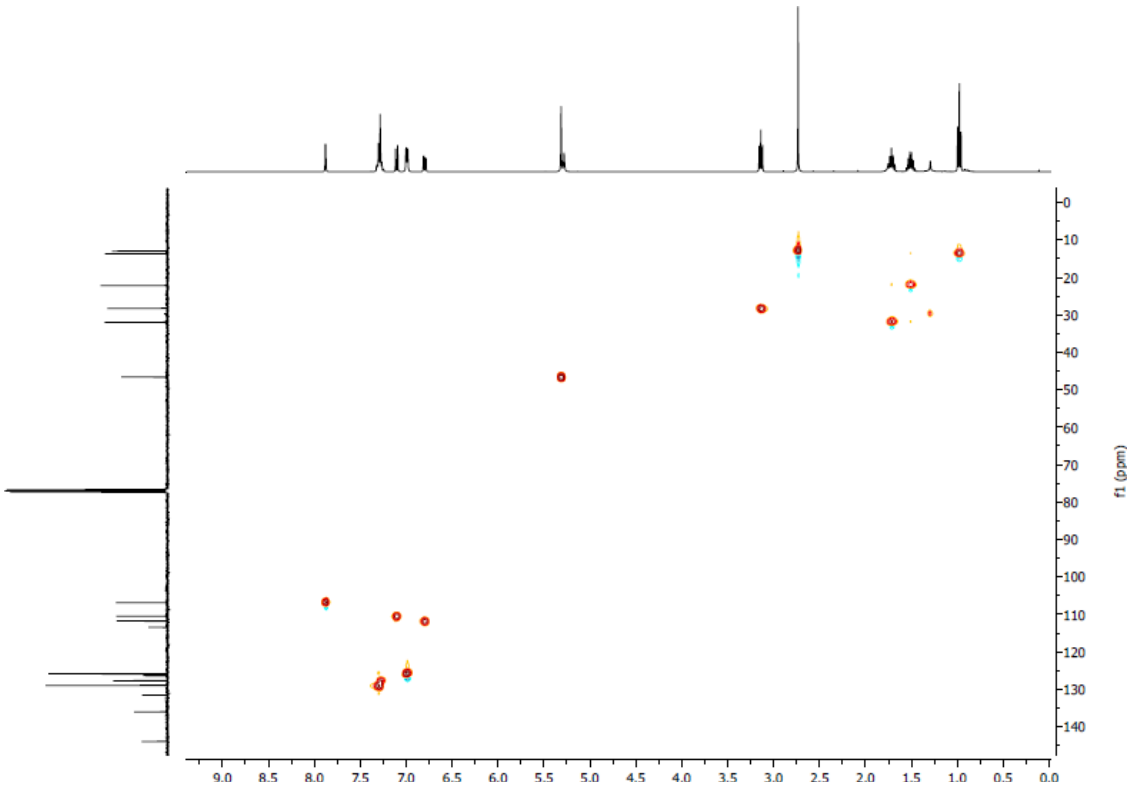
DEPT135



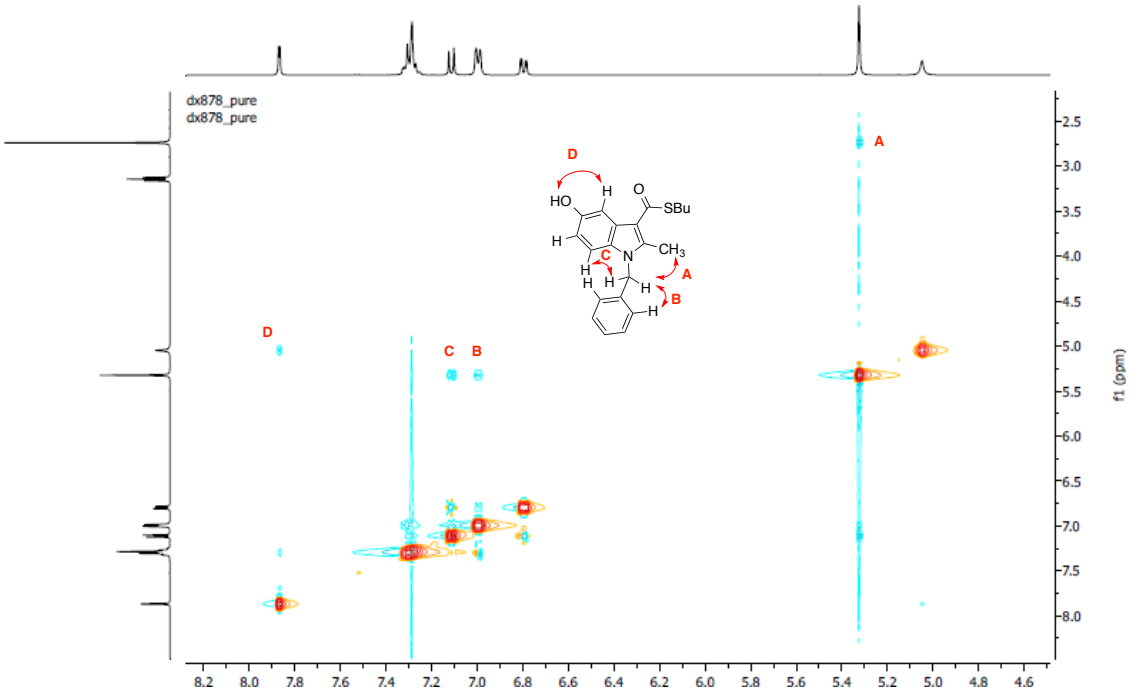
^{1}H-^{1}H COSY



^{1}H-^{13}C HSQC

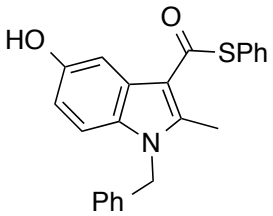


NOESY (Strong NOEs are labeled **A-D**)



Compound **22b**

S-Phenyl 1-benzyl-5-hydroxy-2-methyl-1*H*-indole-3-carbothioate

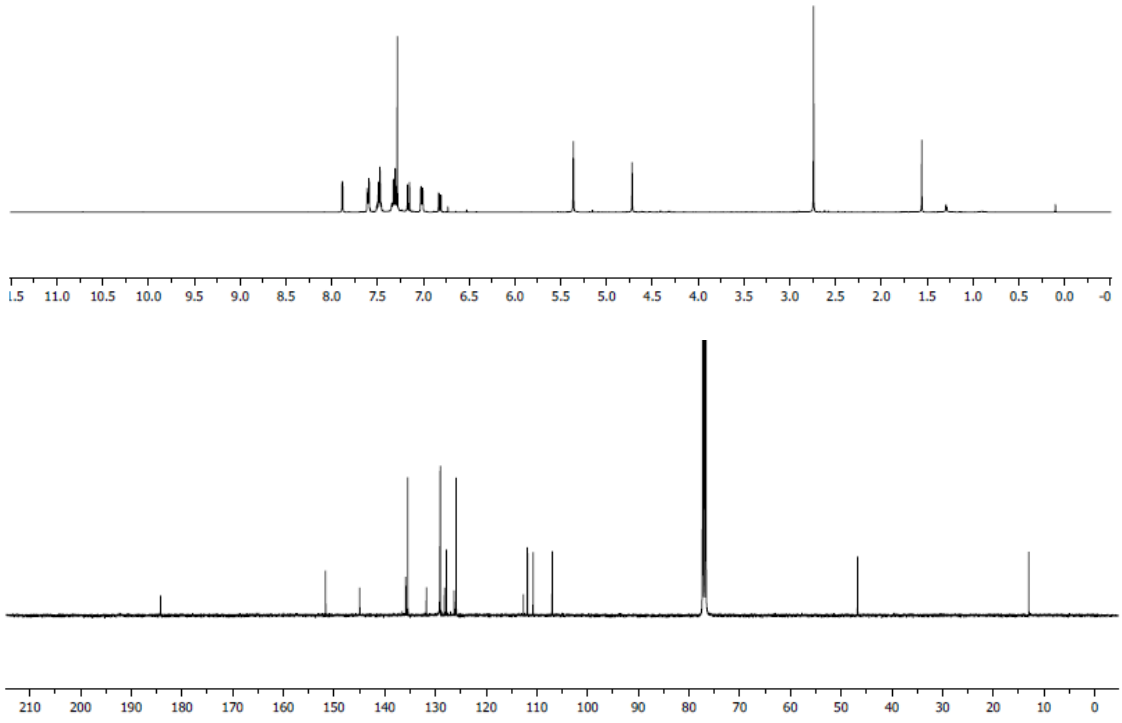


White solid, 68 %; mp = 176.6-177.5 °C;

^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.88 (d, *J* = 2.4 Hz, 1H), 7.65-7.57 (m, 2H), 7.51-7.45 (m, 3H), 7.40-7.26 (m, 3H), 7.16 (d, *J* = 8.7 Hz, 1H), 7.06-6.98 (m, 2H), 6.82 (dd, *J* = 8.8, 2.4 Hz, 1H), 5.36 (s, 2H), 4.72 (s, 1H), 2.74 (s, 3H);

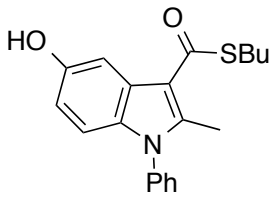
^{13}C-NMR (100 MHz, $CDCl_3$) δ 184.2, 151.6, 144.9, 135.8, 135.4, 131.7, 129.1, 129.0, 129.0, 128.2, 127.8, 126.3, 125.9, 112.6, 111.8, 110.7, 106.9, 46.7, 12.9;

HRMS (MALDI+) m/z calcd for $C_{23}H_{19}NO_2S$ $(M+Li)^+$ 380.1297; found 380.1278.



Compound **22c**

S-Butyl 5-hydroxy-2-methyl-1-phenyl-1*H*-indole-3-carbothioate

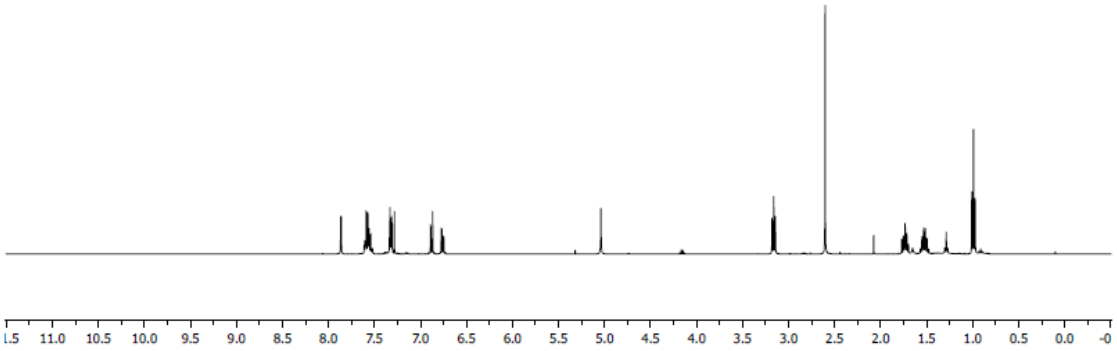


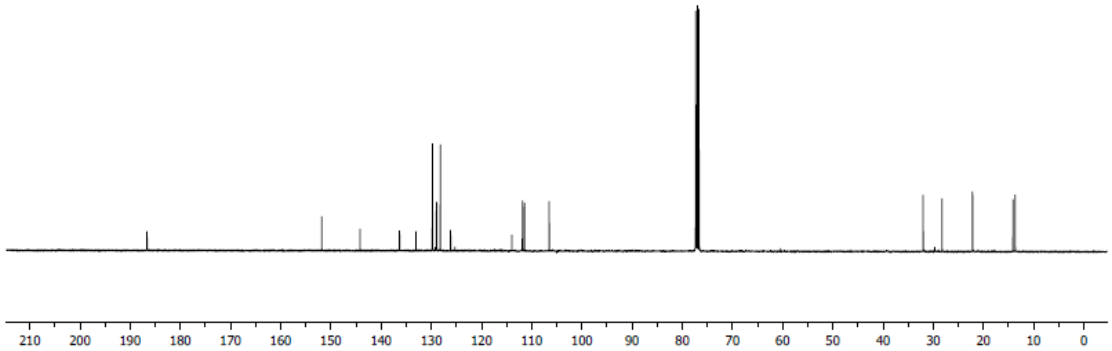
Light yellow solid, 53 %; mp = 180.2-181.9 °C;

^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.86 (d, *J* = 2.3 Hz, 1H), 7.63-7.51 (m, 3H), 7.36-7.27 (m, 2H), 6.88 (d, *J* = 8.7 Hz, 1H), 6.76 (dd, *J* = 8.7, 2.4 Hz, 1H), 5.04 (s, 1H), 3.16 (t, *J* = 7.3 Hz, 2H), 2.60 (s, 3H), 1.80-1.67 (m, 2H), 1.56-1.48 (m, 2H), 0.99 (t, *J* = 7.3 Hz, 3H);

^{13}C-NMR (100 MHz, $CDCl_3$) δ 186.6, 151.8, 144.1, 136.3, 133.0, 129.7, 128.9, 128.2, 126.1, 113.9, 111.8, 111.4, 106.5, 32.0, 28.2, 22.1, 14.1, 13.6;

HRMS (MALDI+) m/z calcd for $C_{20}H_{21}NO_2S$ $(M+H)^+$ 340.1371; found 340.1366.



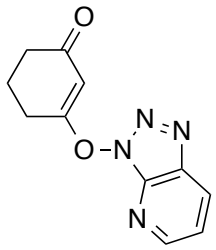


E. Intermediate Formation in HOAt-mediated Reactions

Method A for the synthesis of β-enamino esters was used in the absence of amine nucleophile. The crude product was purified by flash chromatography (CH_2Cl_2 to 5 % EtOAc in CH_2Cl_2) to give compound **23**.

Compound **23**

3-((3*H*-[1,2,3]Triazolo[4,5-*b*]pyridin-3-yl)oxy)cyclohex-2-en-1-one

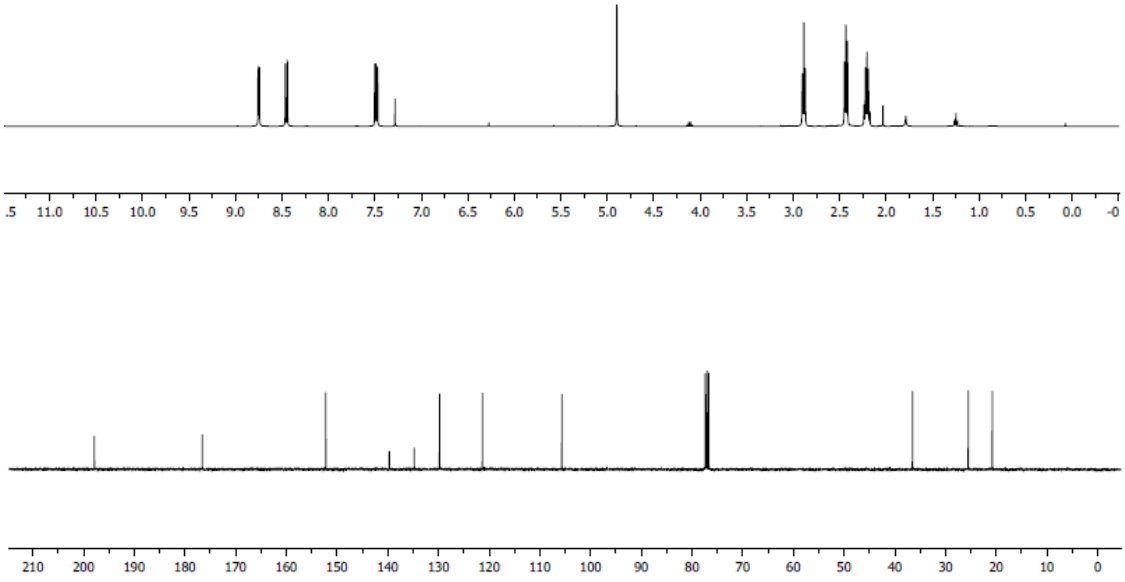


White solid, 65 %; mp = 108.1-109.0 °C;

^{1}H-NMR (400 MHz, $CDCl_3$) δ 8.75 (dd, *J* = 4.5, 1.4 Hz, 1H), 8.45 (dd, *J* = 8.4, 1.4 Hz, 1H), 7.49 (dd, *J* = 8.4, 4.5 Hz, 1H), 4.90 (s, 1H), 2.89 (t, *J* = 6.3, 2H), 2.43 (dd, *J* = 7.4, 6.0 Hz, 2H), 2.25-2.17 (m, 2H);

^{13}C-NMR (100 MHz, $CDCl_3$) δ ^{13}C NMR (101 MHz, $CDCl_3$) δ 197.8, 176.5, 152.2, 139.6, 134.7, 129.7, 121.3, 105.6, 36.5, 25.5, 20.7;

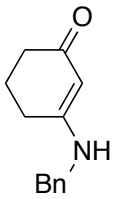
HRMS (ESI+) m/z calcd for $C_{11}H_{10}N_4O_2$ $(M+H)^+$ 231.0882; found 231.0883.



23 (23 mg, 0.1 mmol) was mixed with benzyl amine (22 uL, 0.2 mmol) in $CHCl_3$ at room temperature. The resulting solution was stirred at room temperature for 18h. The solvent was removed and the product was purified by flash chromatography (3 - 6 % MeOH in CH_2Cl_2) to give compound **24**.

Compound **24**

3-(Benzylamino)cyclohex-2-en-1-one

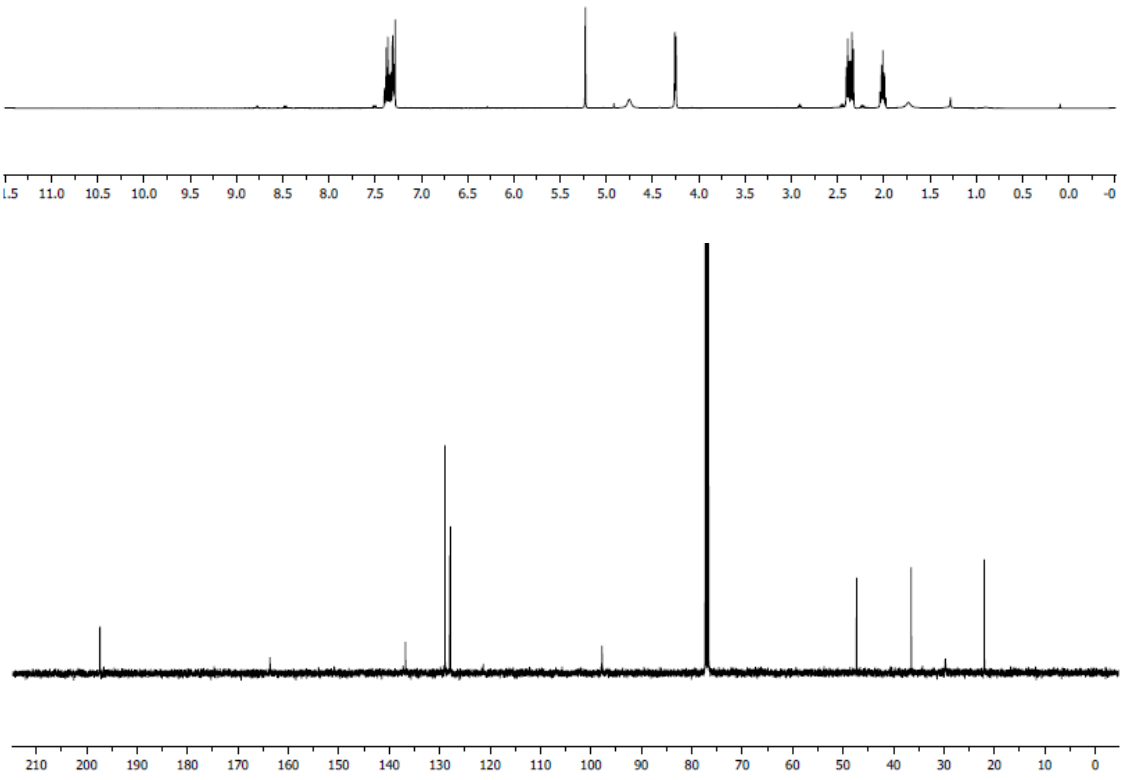


Light yellow oil, 71 %;

^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.43-7.26 (m, 5H), 5.23 (s, 1H), 4.75 (s, 1H), 4.25 (d, *J* = 5.2 Hz, 2H), 2.44-2.30 (m, 4H), 2.05-1.97 (m, 2H);

^{13}C-NMR (100 MHz, $CDCl_3$) δ 197.3, 163.6, 136.7, 128.9, 128.0, 127.8, 97.7, 47.3, 36.4, 29.6, 21.9;

HRMS (ESI+) m/z calcd for $C_{13}H_{15}NO$ $(M+H)^+$ 202.1232; found 202.1239.



F. Syntheses of β-Enamino Amides

Method A was used for the synthesis of β-enamino amides and the crude product was purified by flash chromatography (3 - 6 % MeOH in CH_2Cl_2). Characterization data for compounds **25a** and **25b** are consistent with previous reported data. [12,13]

APPENDIX D

EXPERIMENTAL PROCEDURES FOR CHAPTER IV

A. Nondenaturing PAGE

General procedure: α-Antithrombin was purchased from Haematologic Technologies Inc. and diluted with pH 7.4 Tris-HCl buffer containing 50 mM tris, 50 mM NaCl, 1 mM EDTA. 20 mM stock solutions of the target compounds in DMSO were used in all the gel experiments.

0.25 mg / mL α-Antithrombin was incubated with 200 fold the target compounds or controls at 50 °C for 1 h. The solutions were put on ice to quench the oligomerization after incubation. Aliquots of the solutions were taken and analyzed by 8 % (w/v) acrylamide nondenaturing PAGE to evaluate the extent of oligomerization.[14] Silver stain was performed according to a literature procedure to visualize the results.[15]

B. α-Antithrombin Oligomerizartion Kinetics Determined By Nondenaturing PAGE

Kinetics of a-antithrombin oligomerization was followed by nondenaturing PAGE according to the literature procedure except that silver staining was used to visualize the protein.[16,17] Since silver staining method generally has a limited linear dynamic range, silver staining for a-antithrombin at various sample loading was performed to calibrate the band intensity relative to the amount of protein (**Figure S1**). The linear range was determined to be between 0 μg and 0.75 μg for protein loading on

gel. The protein loading in the kinetics studies was kept in this range to make sure the linear correlation between the amount of monomer and band intensity after silver staining.

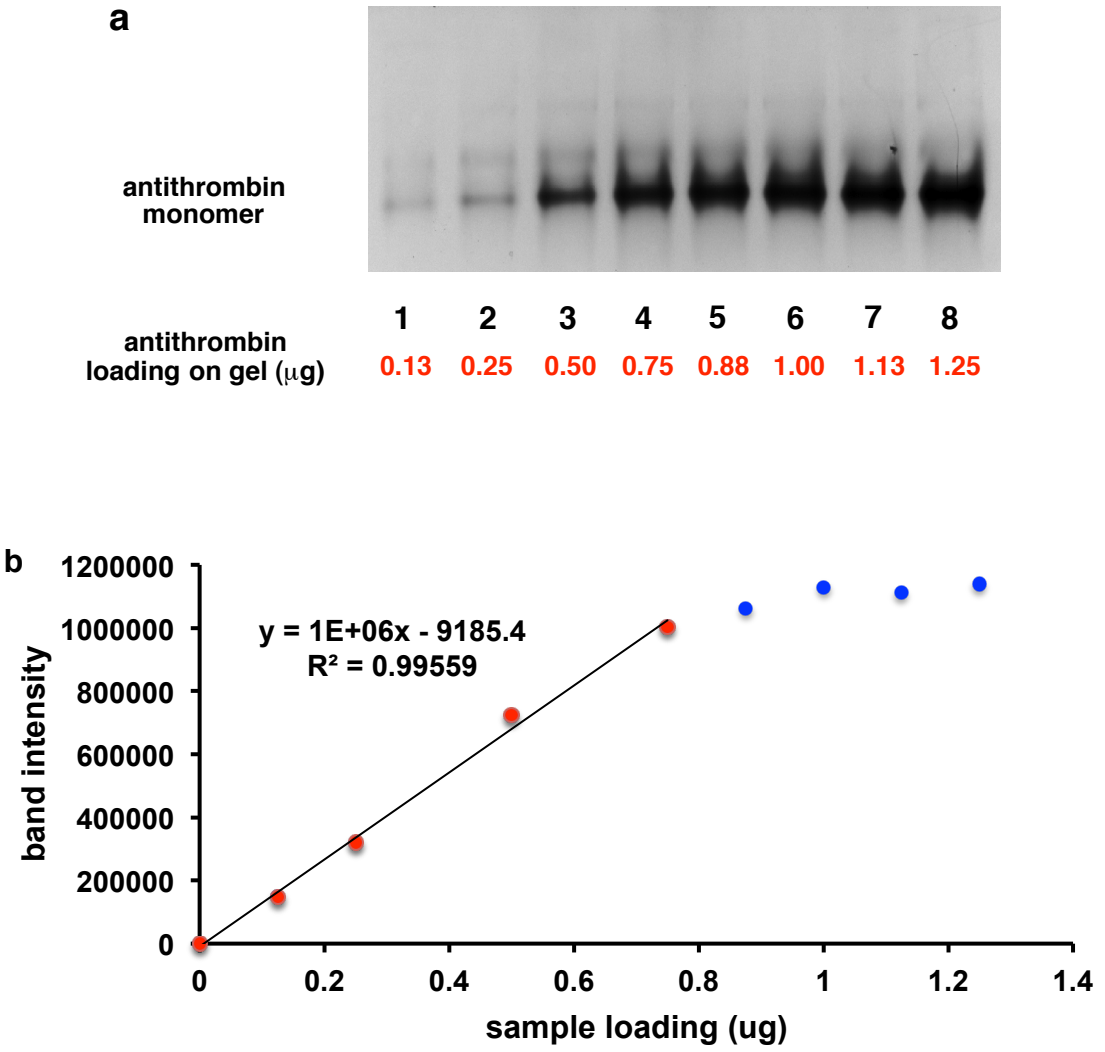
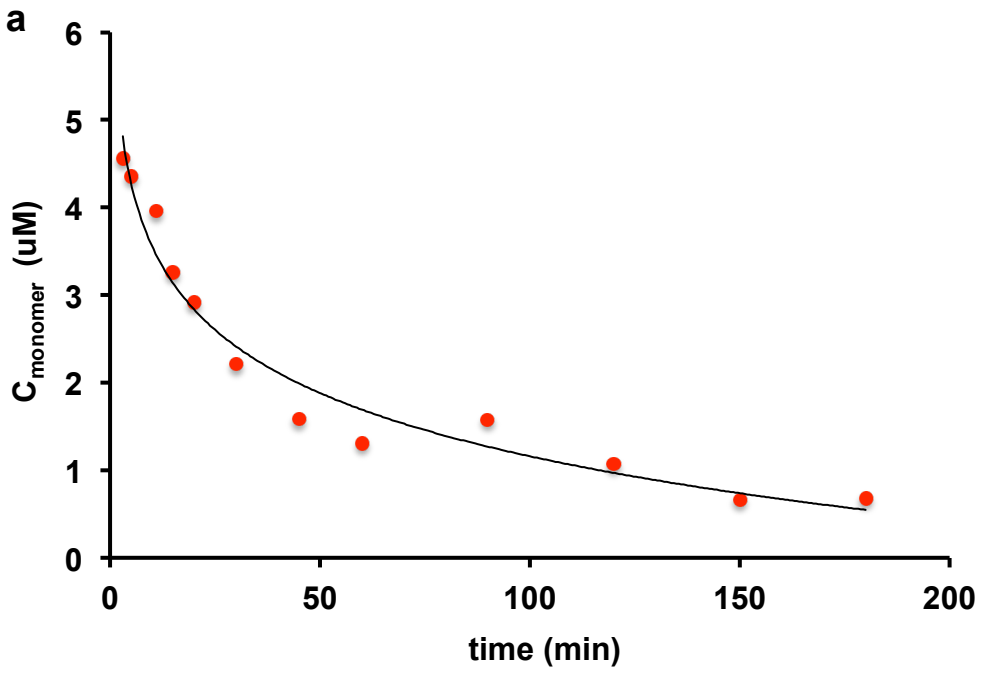


Figure S1. a Band intensity of α-antithrombin monomer at different protein loading on nondenaturing PAGE; b Determination of the linear dynamic range for silver staining.

General procedure: 0.25 mg / mL α-Antithrombin was incubated with 200 fold DDD-**1as'f** at 50 °C for 3 h. Aliquots of the solutions were taken and cold quenched on ice in every 10 mins in the first hour and then in every 30 min afterwards. The relative amount of residual monomer was analyzed by 8 % (w/v) acrylamide nondenaturing PAGE. The intensities of the a-antithrombin monomer bands were measured and fitted to a second order kinetics model to determine the rate constant in terms of the loss of monomer. A representative example for DDD-**1as'f** was shown in Figure S2.



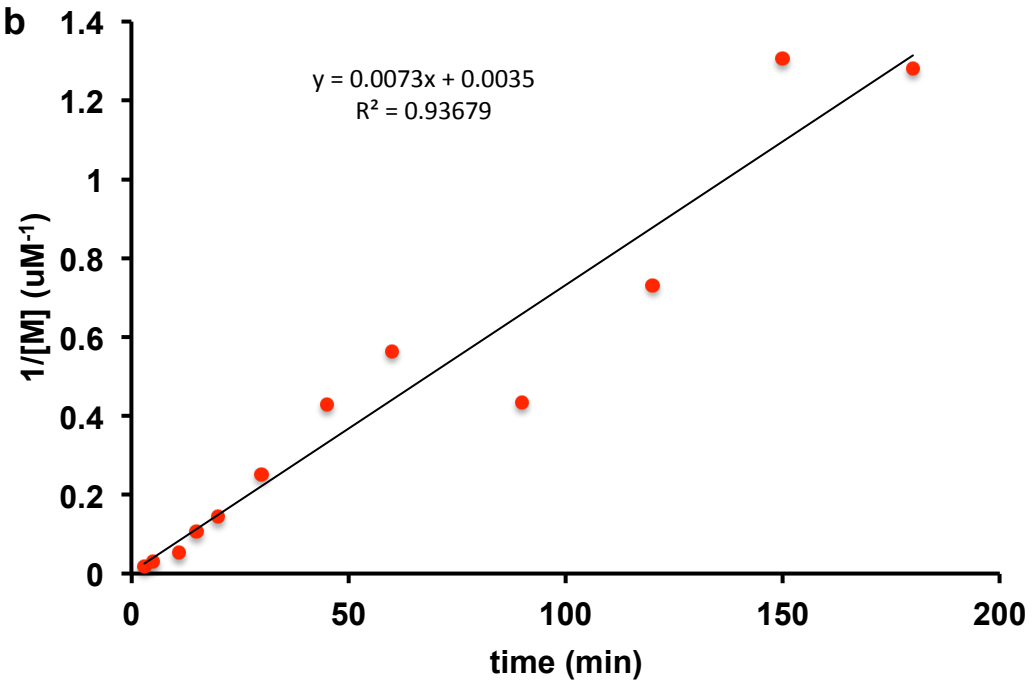


Figure S2. a Loss of monomeric α-antithrombin measured by nondenaturing PAGE; b Fitting into a second order kinetics model.

C. Thrombin Activity Assay

0.25 mg / mL a-Antithrombin was incubated with 200 fold DDD-**1as'f** or DDD-**1asf** at 50 °C for 3 -5 h. Aliquots of the solutions were taken at indicated time and quenched on ice to stop oligomerization. The antithrombin solutions at different degrees of oligomerization were then incubated with 4 nM thrombin (purchased from Haematologic Technologies Inc.) in 20 mM Tris-HCl buffer containing 150 mM NaCl, 2.5 mM $CaCl_2$, 0.1 % PEG8000, 0.02 % Tween80 at pH = 7.4 in the presence of 10 nM heparin. After incubation of 5 min at room temperature, chromogenic substrate Spectrozyme TH (purchased from American diagnostic inc.) was added. The residual

enzyme activity was determined from the initial rate of increase in absorbance at 405 nm with a plate reader.[18]

D. Electron Microscopy

0.25 mg / mL α-Antithrombin was incubated with or without 200 fold DDD-**1as'f** at 50 °C for 1 h. The solutions were put on ice to quench the oligomerization after incubation. The solution was diluted to a concentration of 0.05 mg / mL in pH 7.4 Tris-HCl buffer containing 50 mM tris, 50 mM NaCl, 1 mM EDTA for use in the EM experiments. Specimens were prepared using an aqueous solution of 1% wt/vol uranyl acetate (pH 4.25). Uranyl acetate has been shown to fix the structure of protein molecules within 10 ms and prior to blotting the buffer off.[19] Therefore it can be reasonably assumed that the preparation reflects the situation in solution. Specimens were observed in a Jeol 1200 EX TEM operated at an accelerating voltage of 100 kV. Images were captured at calibrated magnifications using an optically coupled 3 k slow scanCCD camera (model 15C, SIA, Duluth, GA) and Maxim DL imaging software.

APPENDIX E

EXPERIMENTAL PROCEDURES FOR CHAPTER V

A. Synthesis of Cyclic Peptides

General Procedures

Coupling Method I

Boc or Cbz protected amino acid (Pg-AA-OH, 2.0 mmol, 1.0 equiv.) was added to a solution of HOAt (408 mg, 3.0 mmol, 1.5 equiv.) and NMM (220 μL, 2.0 mmol, 1.0 equiv.) in DCM (3.0 mL) at 0 ^{o}C under N_2. The resulting mixture was stirred at 0 ^{o}C for 10 min and then methyl anthranilate (906 mg, 6.0 mmol, 3.0 equiv.) was added, followed by the addition of EDC•HCl (575 mg, 3.0 mmol, 1.5 equiv.) in one portion. The reaction mixture was allowed to warm to room temperature over 1 h and was stirred at room temperature for 6 h. 30 mL DCM was added to dilute the solution and the organic phase was washed with 0.2 M HCl aqueous solution (30 mL × 5). The organic phase was further washed with brine (10 mL), saturated $NaHCO_3$ solution (20 mL × 2) and brine (10 mL), dried over $MgSO_4$ and filtered. The solvent was removed under vacuum to give the crude material. The crude material was used in the next step without further purification. An analytically pure sample was prepared by crystallization from DCM/hexanes.

Coupling Method II

EDC•HCl (287 mg, 1.5 mmol, 1.5 equiv.) was added to a mixture of Boc or Cbz protected amino acid (Pg-AA-OH, 1.0 mmol, 1.0 equiv.), *N*-terminus deprotected linear

peptide (1.0 mmol, 1.0 equiv.), HOBt (203 mg, 1.5 mmol, 1.5 equiv.) and DIPEA (175 μL, 1.0 mmol, 1.0 equiv.) in DCM (10.0 mL) at 0 ^{o}C under N_2. The reaction mixture was stirred at 0 ^{o}C for 10 min and was allowed to warm to room temperature. After overnight reaction at room temperature, 40 mL DCM was added to dilute the solution and the organic phase was washed with 0.1 M HCl aqueous solution (30 mL × 2). The organic phase was further washed with brine (10 mL), saturated $NaHCO_3$ solution (30 mL × 2) and brine (10 mL), dried over $MgSO_4$ and filtered. The solvent was removed under vacuum to give the crude material. The crude material was used in the next step without further purification. An analytically pure sample can be prepared by crystallization from DCM/hexanes.

Boc Deprotection Procedure I

Boc protected intermediate (1.0 mmol) was dissolved in DCM (5 mL) at 0 ^{o}C, and the mixture was stirred at 0 ^{o}C for 10 min. TFA (5 mL) was added in one portion and the reaction mixture was stirred for 30 min at room temperature. Toluene (30 mL) was added and the solution was concentrated. Residual TFA was azeotroped 3 times with toluene (3 × 30 mL) to give the crude product. Saturated $NaHCO_3$ (20 mL) was added and the aqueous layer was extracted with DCM (30 mL). The organic layer was further washed with saturated $NaHCO_3$ (20 mL × 2) and brine (10 mL), dried over $MgSO_4$ and filtered. The solvent was removed under vacuum to give the crude material. The crude material was used in the next step without further purification.

Boc Deprotection Procedure II

4 M HCl in dioxane (5 mL) was added to Boc protected intermediate (1.0 mmol) at 0 ^{o}C, and the mixture was stirred at room temperature for 30 min. Toluene (20 mL) was added and the solution was concentrated. Residual dioxane was azeotroped 3 times with toluene (3 × 20 mL) and then dried under high vacuum for 3 h to give the HCl salt of the Boc deprotected product.

^{t}Bu Deprotection Procedure

^{t}Bu protected cyclic peptides (0.08 mmol) was dissolved in 1:1 TFA/DCM containing 5 % (v/v) triethylsilane (TES) (0.8 mL) at 0 ^{o}C, and the mixture was stirred at 0 ^{o}C for 10 min. The reaction mixture was stirred for an additional 60 min at room temperature. Toluene (2 mL) was added and the solution was concentrated. Residual TFA was azeotroped 3 times with toluene (3 × 2 mL) to give the crude product. The crude product was purified with flash chromatography (4 – 8 % MeOH in DCM containing 0.1 % AcOH) to give the pure product.

Cbz Deprotection Procedure

To a solution of Cbz protected substrate in methanol (0.1 M) under nitrogen was added 10 wt % Pd/C (0.05 equiv. Pd). The reaction was placed under an atmosphere of hydrogen (1 atm, balloon) for 12 h. After the reaction finished, the flask was purged with N_2. The reaction mixture was filtered over a Celite pad and concentrated to afford the product. The product was used in the next step without further purification.

Bn Deprotection Procedure

To a solution of Cbz protected substrate in methanol (0.05 M) under nitrogen was added 10 wt % Pd/C (0.15 equiv. Pd). The reaction was placed under an atmosphere of hydrogen (1 atm, balloon) for 24 h. After the reaction finished, the flask was purged with N_2. The reaction mixture was filtered over a Celite pad and concentrated to afford the product. The product was purified with flash chromatography (4 – 8 % MeOH in CH_2Cl_2) to give the pure product.

Hydrolysis of Methyl Ester

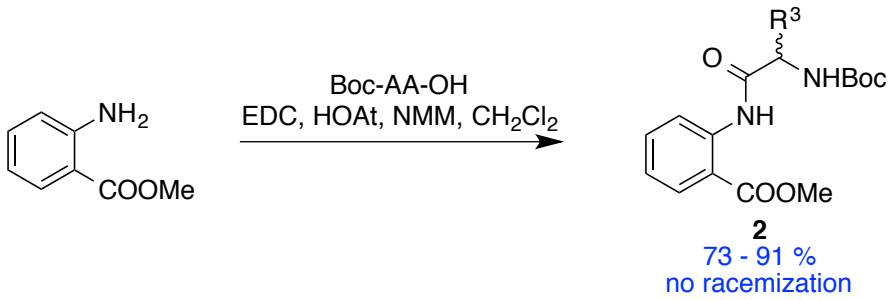
The linear tetrapeptide (0.2 mmol) was added to 2.7 mL THF and the resulting suspension was stirred at room temperature for 10 min. A 0.3 M aqueous solution of LiOH (1.3 mL, 0.4 mmol, 2.0 equiv.) was added and the resulting mixture was stirred at room temperature for 2h. The mixture was concentrated to remove most THF and 0.1 M HCl solution (10 mL) was added. The white solid was filtered, washed with 0.1 M HCl and then water and dried under vacuum to give the product.

Procedure for Cyclization

The deprotected linear peptide (0.2 mmol, 1.0 equiv.) was dissolved in DMF/DCM (1:1 mixture) to give a 2 mM final concentration. HOAt (81.6 mg, 0.6 mmol, 3.0 equiv.) and NMM (88 μL, 0.8 mmol, 4.0 equiv.) were added to the solution, followed by EDC•HCl (114.6 mg, 0.6 mmol, 3.0 equiv.). The resulting mixture was

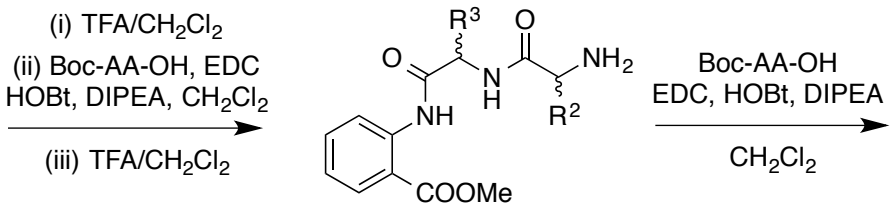
stirred at room temperature under N_2 for 48 h and then concentrated in vacuum to give the crude mixture. The crude product was purified with flash chromatography (3 – 7 % MeOH in DCM) to give the pure product.

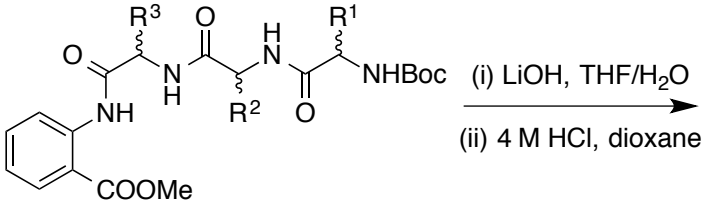
Boc Approach to Products 1



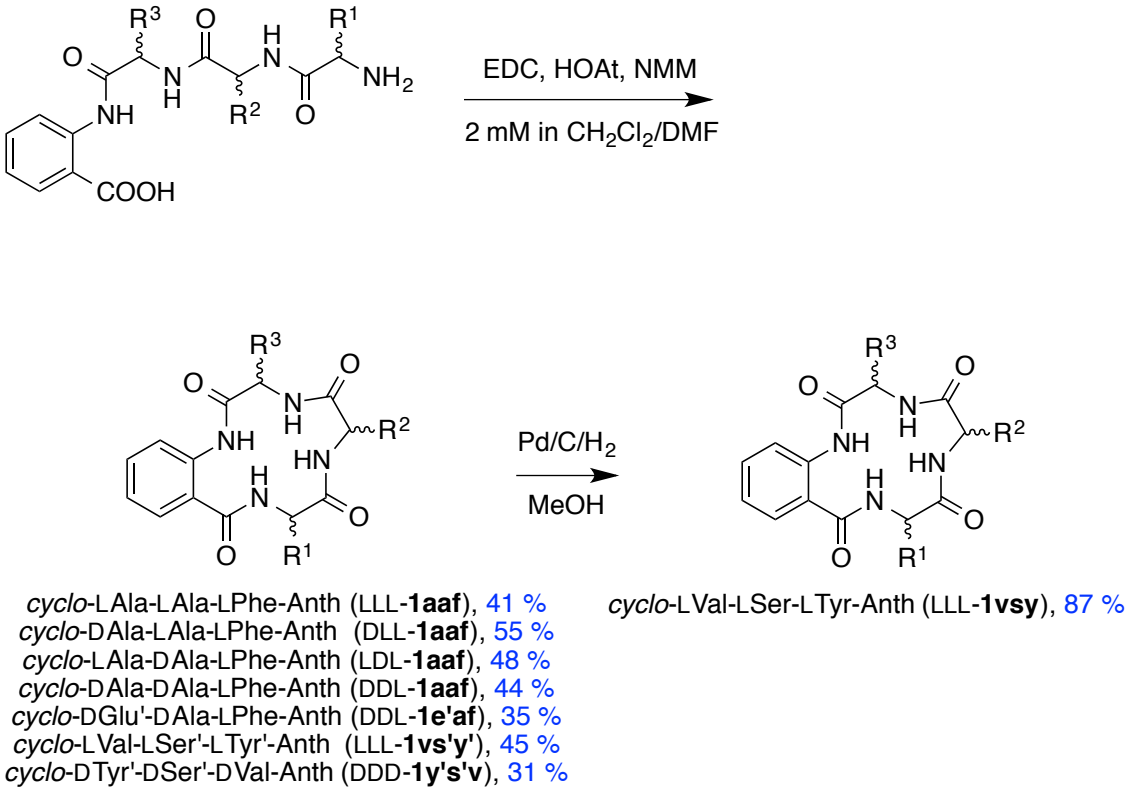
2

73 - 91 %
no racemization





70 - 90 % over four steps



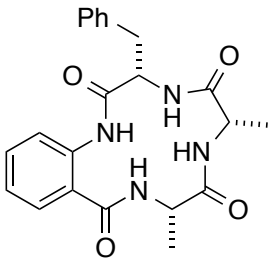
Scheme S1. Boc approach to products **1**

General Procedures for Boc Approach in Solution

Methyl anthranilate was coupled to the first Boc protected amino acid with "Coupling Method I". The Boc protection group was deprotected with procedure described in "Boc Deprotection Procedure I" and the resulting material was coupled with the second Boc protected amino acid with "Coupling Method II". The tripeptide intermediate was deprotected with "Boc Deprotection Procedure I" and coupled with the third Boc amino acid with "Coupling Method II" to give the protected tetrapeptide intermediate. The methyl ester of this intermediate was hydrolyzed with the procedure "Hydrolysis of Methyl Ester" and the N- terminus Boc group was removed with "Boc

Deprotection Procedure II". The linear tetrapeptide was cyclized with "Procedure for Cyclization" to give the cyclic peptide product. If necessary, the cyclic peptide product was deprotected with "Bn Deprotection Procedure" or "^{t}Bu Deprotection Procedure" to give the final deprotected product.

(3S,6S,9S)-3-Benzyl-6,9-dimethyl-3,4,6,7,9,10-hexahydro-1H-benzo[k][1,4,7,10]tetraazacyclotridecine-2,5,8,11-tetraone (LLL-**1aaf**)



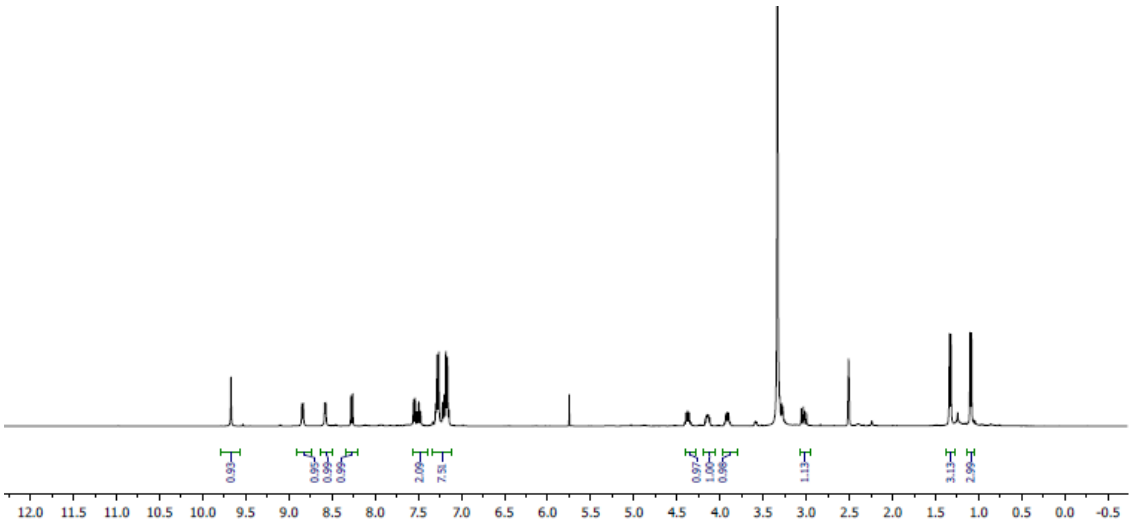
White solid, 41 % over three steps; mp > 270 °C;

^{1}H-NMR (400 MHz, DMSO-d_6) δ 9.67 (s, 1H), 8.84 (d, J = 7.0 Hz, 1H), 8.58 (d, J = 5.0 Hz, 1H), 8.27 (d, J = 7.9 Hz, 1H), 7.57-7.40 (m, 2H), 7.34-7.11 (m, 7H), 4.42-4.33 (m, 1H), 4.19-4.10 (m, 1H), 3.95-3.88 (m, 1H), 3.28-3.35 (m, 1H), 3.03 (dd, J = 13.9, 9.7 Hz, 1H), 1.33 (d, J = 7.4 Hz, 3H), 1.09 (d, J = 6.8 Hz, 3H);

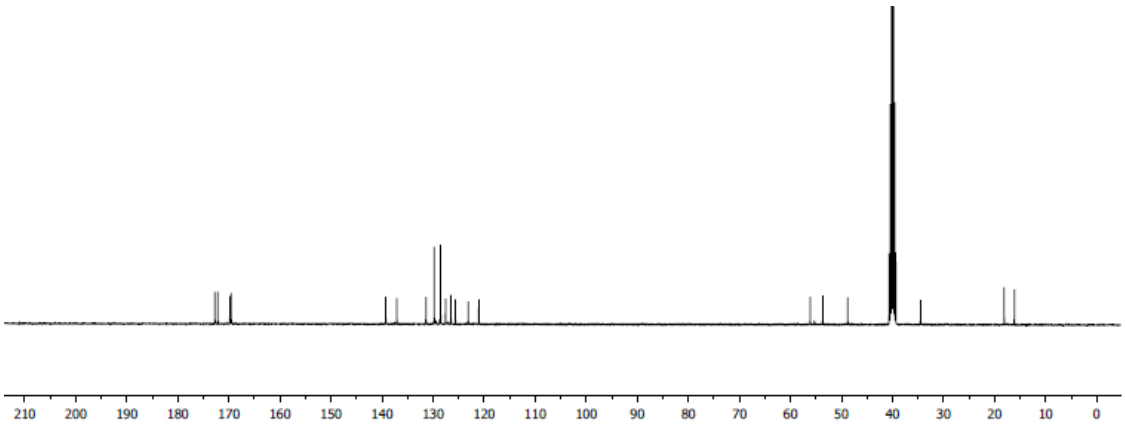
^{13}C-NMR (101 MHz, DMSO-d_6) δ 172.7, 172.1, 169.8, 169.5, 139.2, 137.1, 131.4, 129.7, 128.5, 127.5, 126.5, 125.6, 123.1, 121.0, 56.1, 53.6, 48.7, 34.4, 18.1, 16.1;

HRMS (ESI+) m/z calcd for $C_{22}H_{24}N_4O_4Na$ $(M+Na)^+$ 431.1695; found 431.1679.

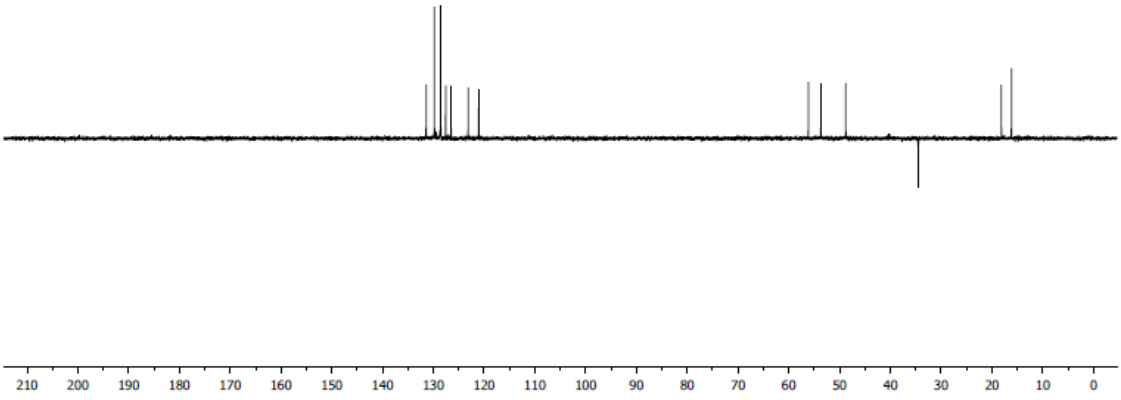
^{1}H NMR



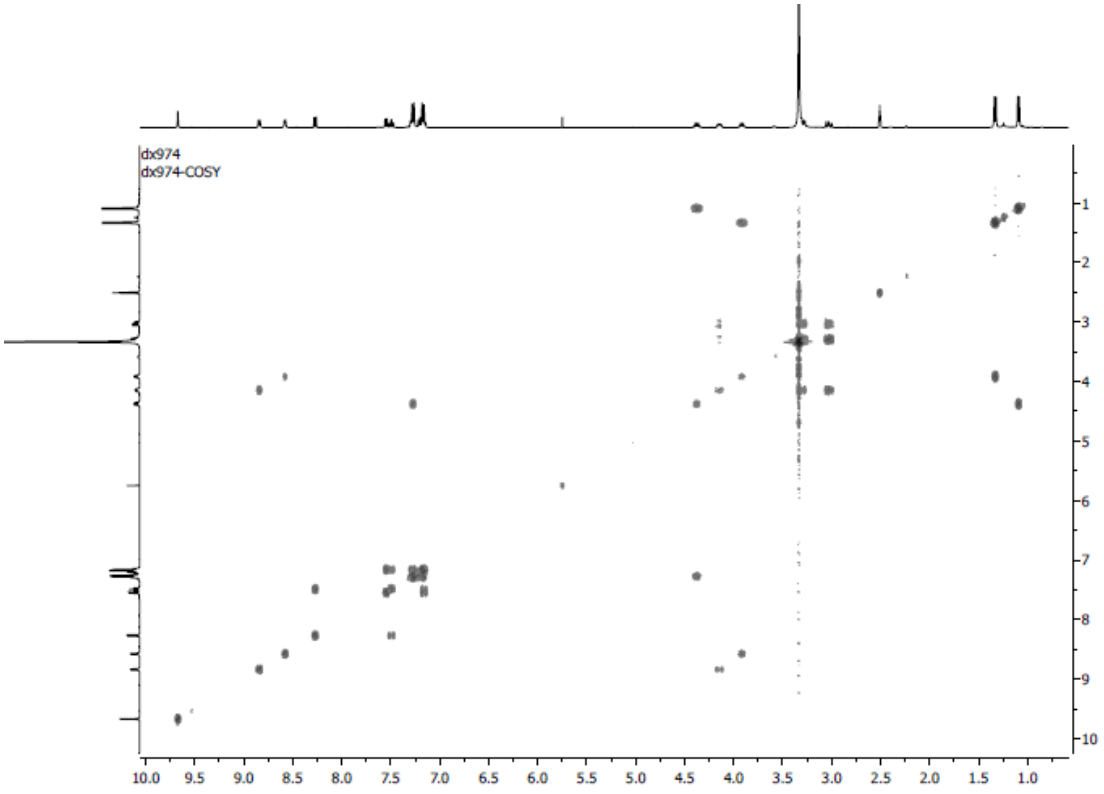
^{13}C NMR



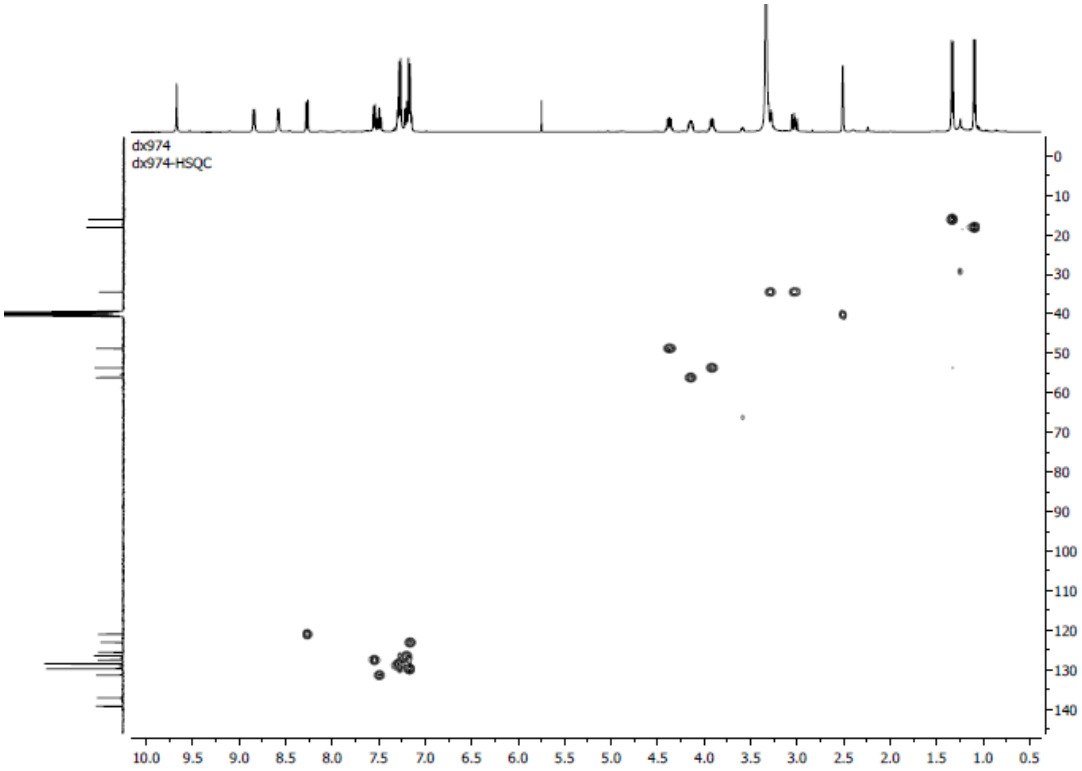
DEPT135



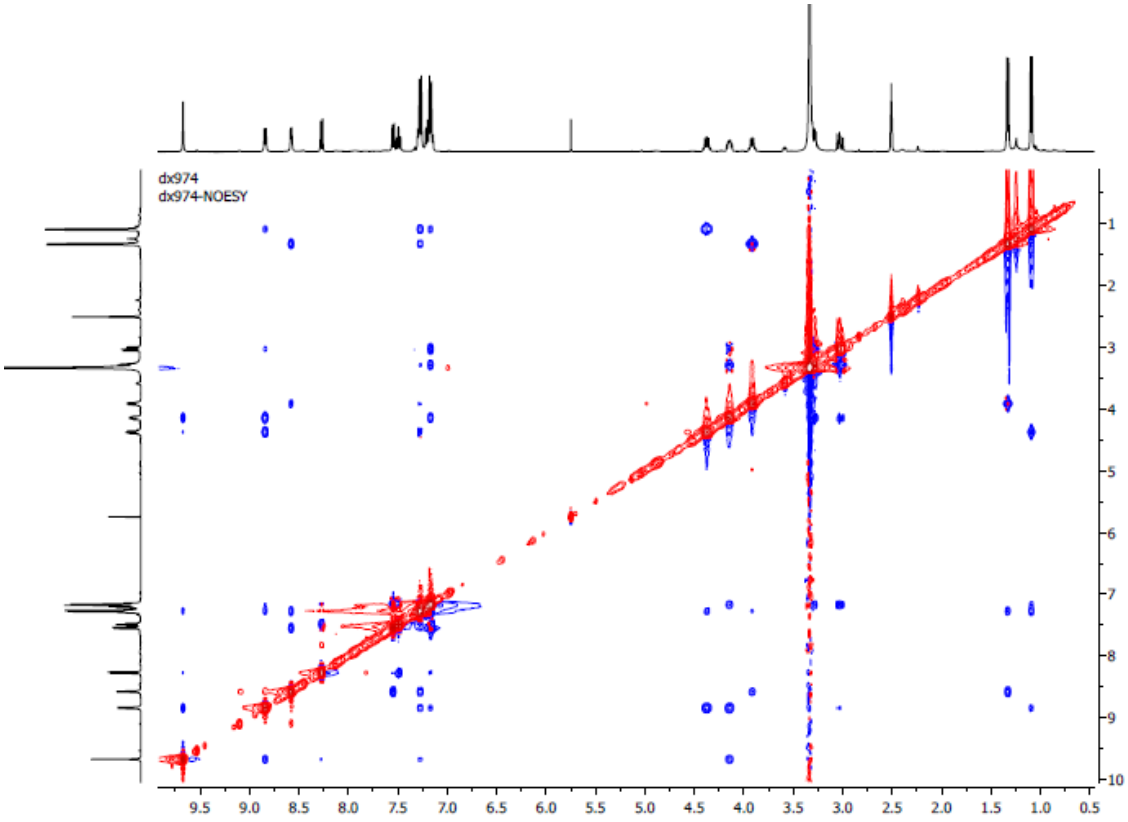
^{1}H-^{1}H COSY



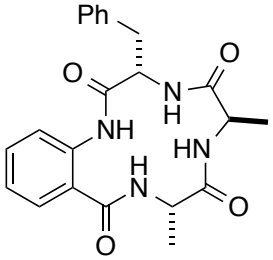
HSQC


dx974
dx974-HSQC
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5

NOESY


dx974
dx974-NOESY
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
1
2
3
4
5
6
7
8
9
10

(3S,6R,9S)-3-Benzyl-6,9-dimethyl-3,4,6,7,9,10-hexahydro-1H-benzo[k][1,4,7,10]tetraazacyclotridecine-2,5,8,11-tetraone (LDL-**1aaf**)



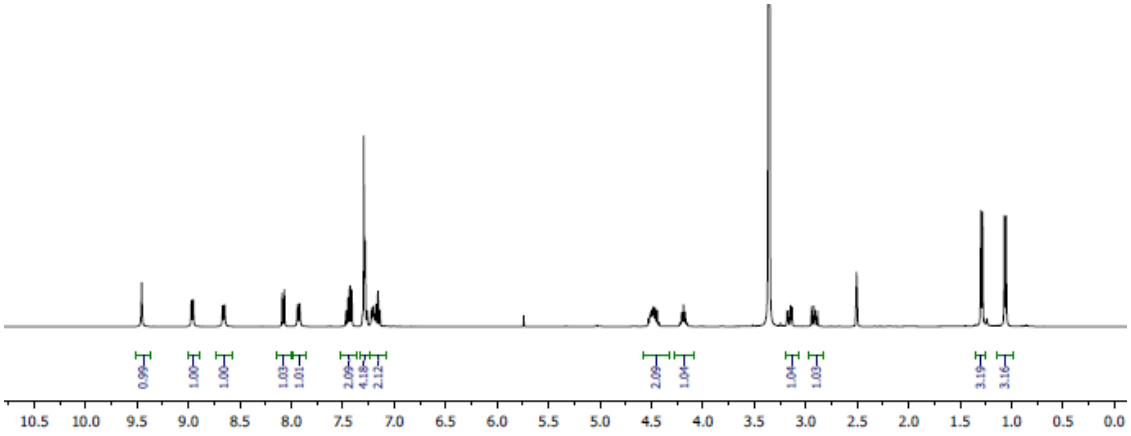
White solid, 48 % over three steps; mp > 270 °C;

^{1}H-NMR (400 MHz, DMSO-d_6) δ 9.45 (s, 1H), 8.96 (d, J = 7.2 Hz, 1H), 8.66 (d, J = 7.7 Hz, 1H), 8.08 (d, J = 8.0 Hz, 1H), 7.93 (d, J = 8.8 Hz, 1H), 7.52-7.36 (m, 2H), 7.33-7.23 (m, 4H), 7.23-7.08 (m, 2H), 4.59-4.32 (m, 2H), 4.23-4.15 (m, 1H), 3.16 (dd, J = 14.0, 5.1 Hz, 1H), 2.91 (dd, J = 14.0, 9.5 Hz, 1H), 1.29 (d, J = 6.9 Hz, 3H), 1.06 (d, J = 6.8 Hz, 3H);

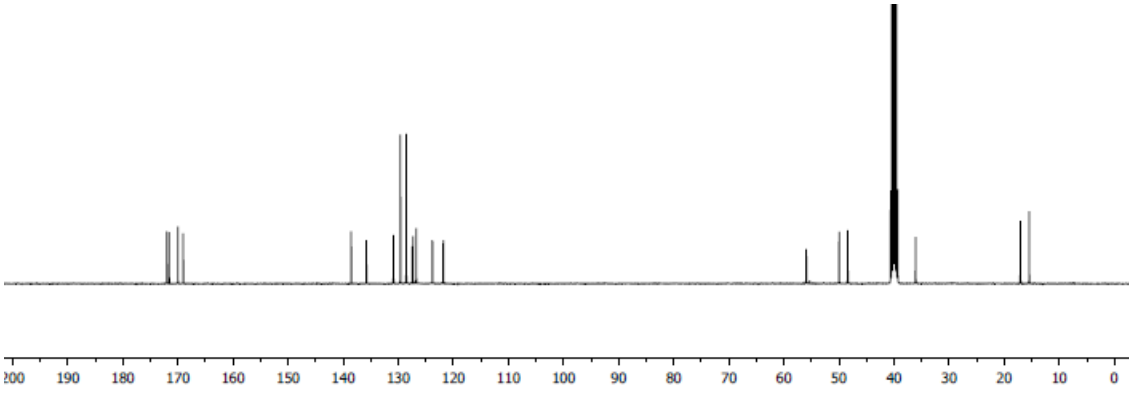
^{13}C-NMR (101 MHz, DMSO-d_6) δ 172.0, 171.5, 170.0, 169.0, 138.5, 135.7, 130.8, 129.6, 128.5, 127.4, 127.3, 126.7, 123.8, 121.8, 55.9, 49.9, 48.4, 36.0, 17.0, 15.4;

HRMS (ESI+) m/z calcd for $C_{22}H_{24}N_4O_4Li$ $(M+Li)^+$ 415.1958; found 415.1948.

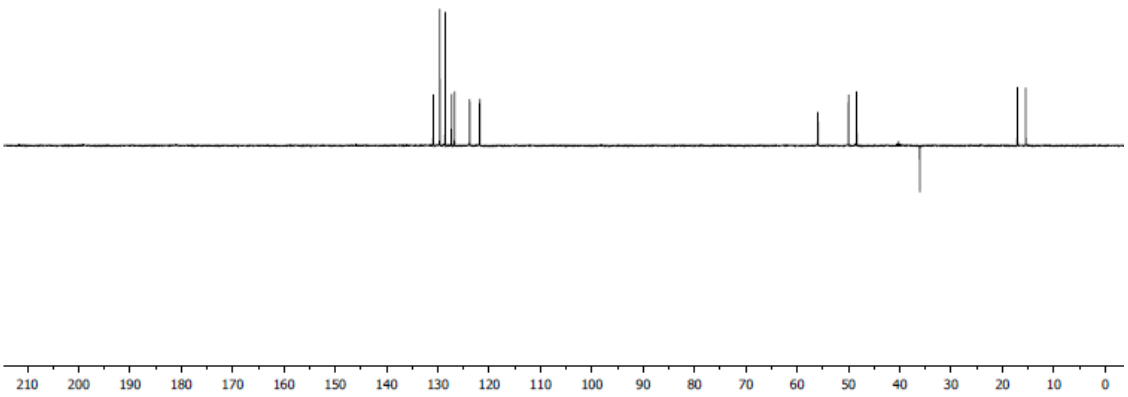
^{1}H NMR



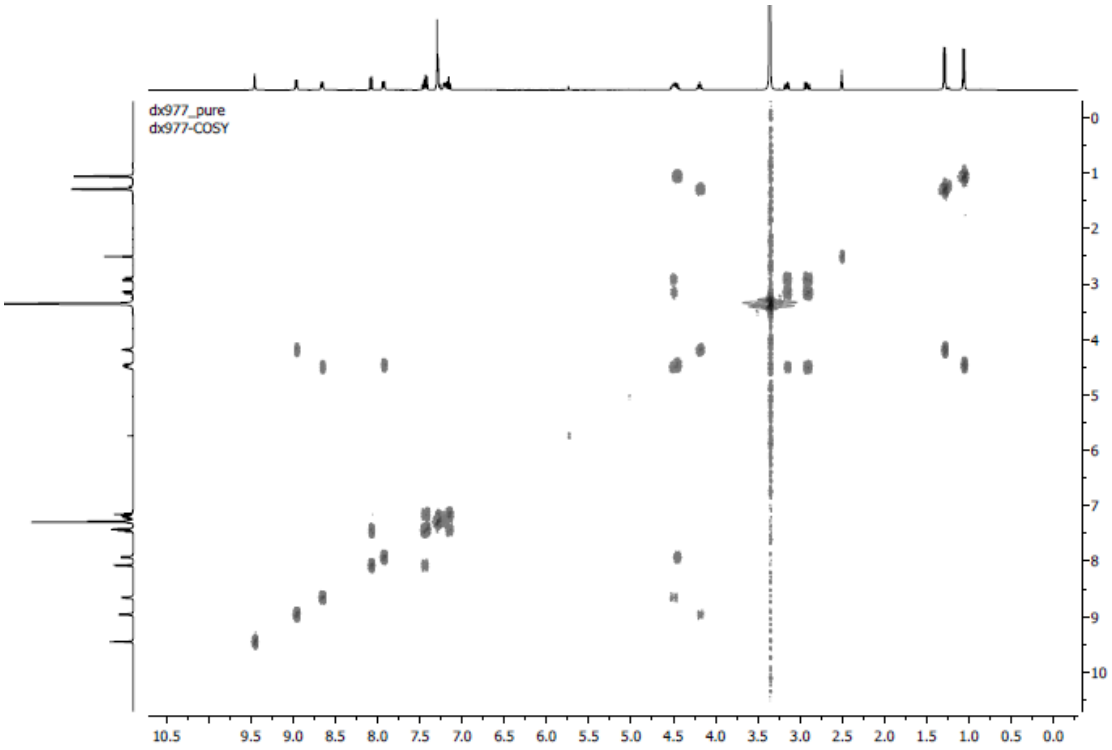
^{13}C NMR



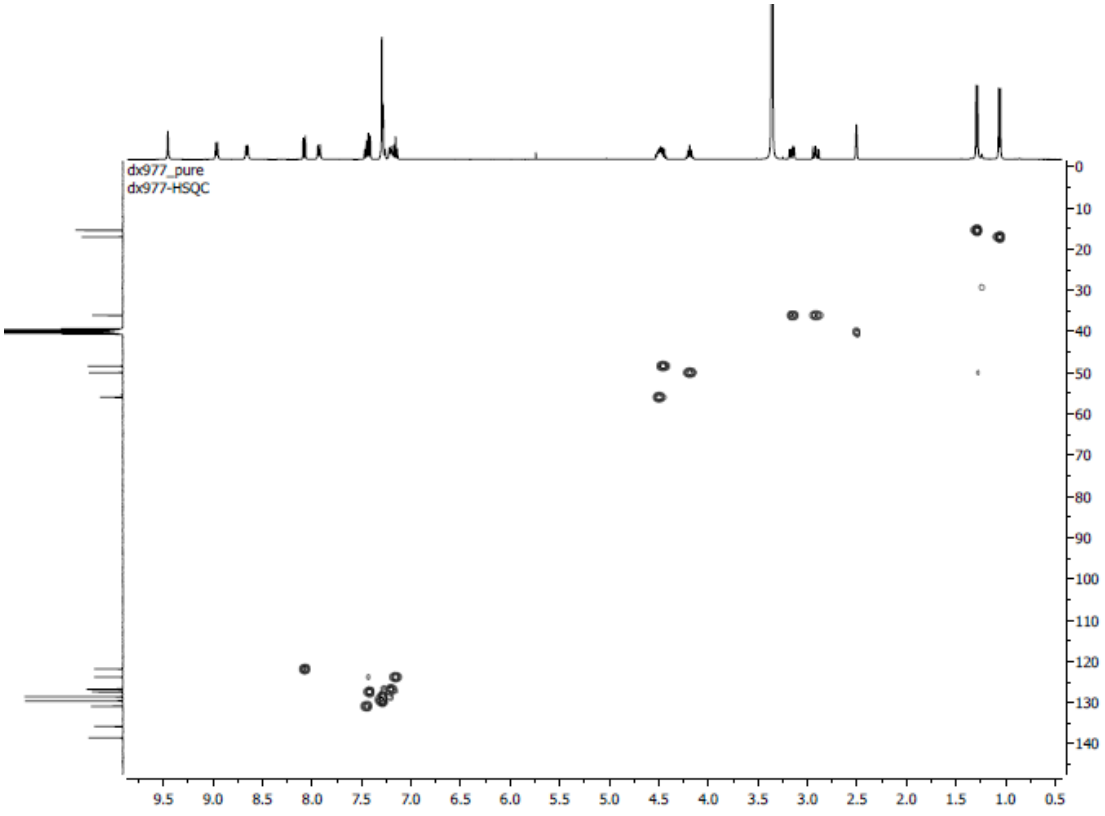
DEPT135


210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0

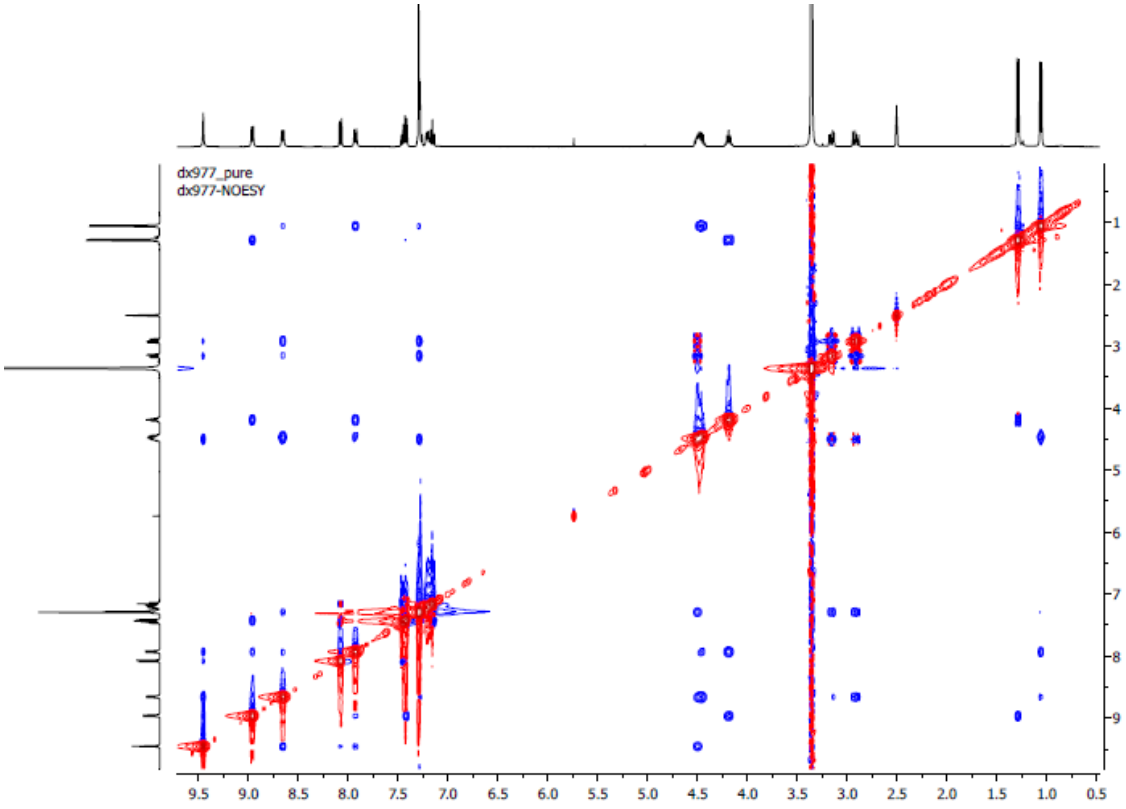
^{1}H-^{1}H COSY



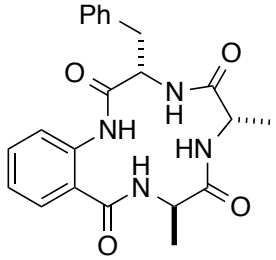
^{1}H-^{13}C HSQC


dx977_pure
dx977-HSQC
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5

NOESY


dx977_pure
dx977-NOESY
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
1
2
3
4
5
6
7
8
9

(3S,6S,9R)-3-Benzyl-6,9-dimethyl-3,4,6,7,9,10-hexahydro-1H-benzo[k][1,4,7,10]tetraazacyclotridecine-2,5,8,11-tetraone (DLL-**1aaf**)



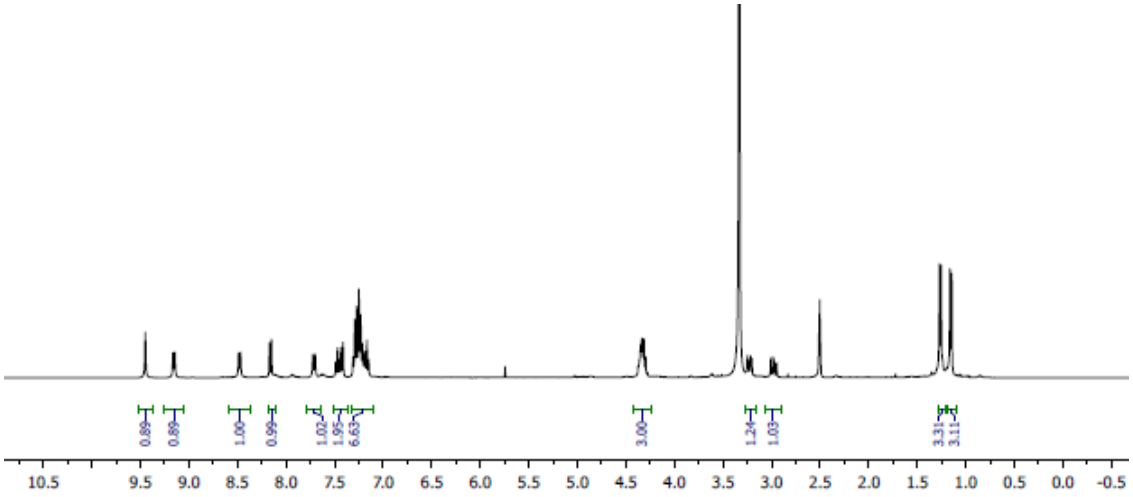
White solid, 55 % over three steps; mp > 270 °C;

^{1}H-NMR (400 MHz, DMSO-d_6) δ 9.45 (s, 1H), 9.15 (d, J = 7.7 Hz, 1H), 8.48 (d, J = 7.9 Hz, 1H), 8.16 (d, J = 8.1 Hz, 1H), 7.71 (d, J = 8.9 Hz, 1H), 7.51-7.36 (m, 2H), 7.33-7.11 (m, 6H), 4.43-4.23 (m, 3H), 3.23 (dd, J = 13.7, 4.5 Hz, 1H), 2.98 (dd, J = 13.8, 10.2 Hz, 1H), 1.26 (d, J = 6.9 Hz, 3H), 1.15 (d, J = 7.2 Hz, 3H);

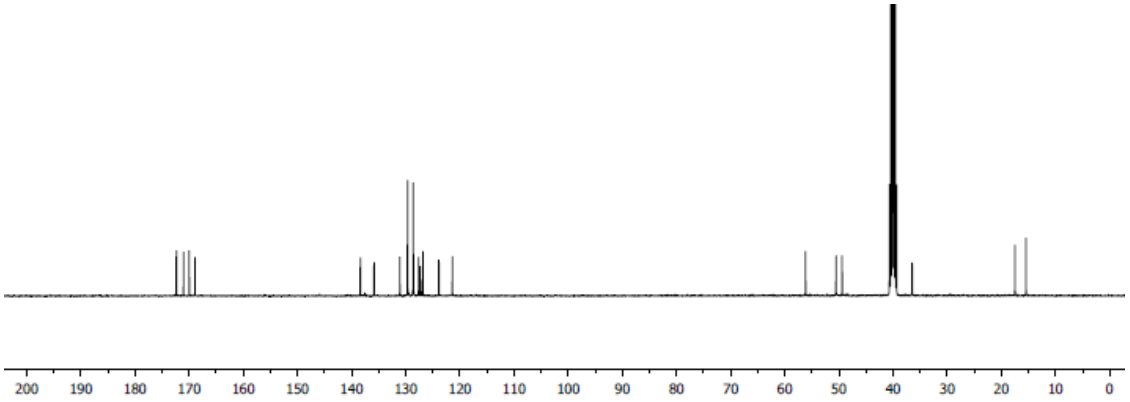
^{13}C-NMR (101 MHz, DMSO-d_6) δ 172.3, 171.0, 169.9, 168.9, 138.3, 135.8, 131.0, 129.6, 128.6, 127.6, 127.3, 126.8, 123.8, 121.3, 56.1, 50.5, 49.4, 36.4, 17.5, 15.4;

HRMS (ESI+) m/z calcd for $C_{22}H_{24}N_4O_4Na$ $(M+Na)^+$ 431.1695; found 431.1682.

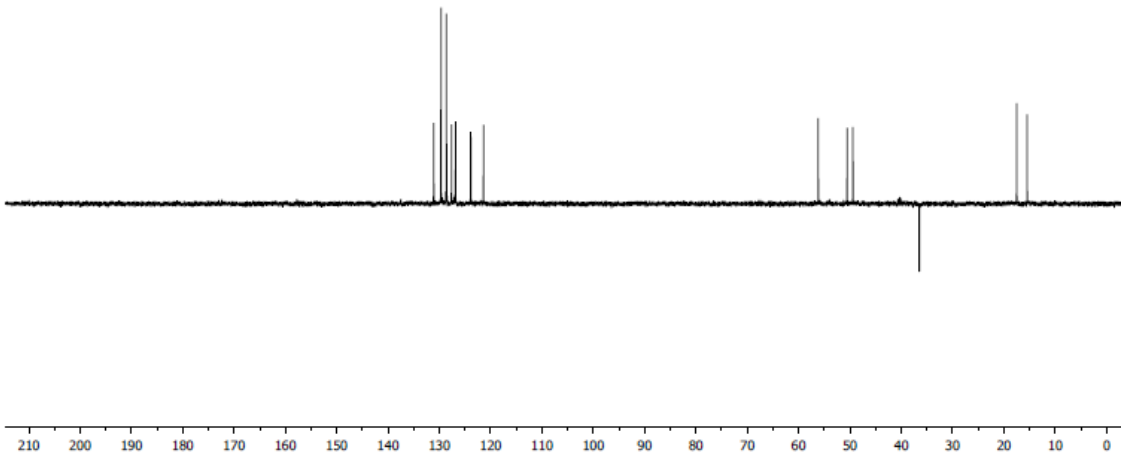
^{1}H NMR



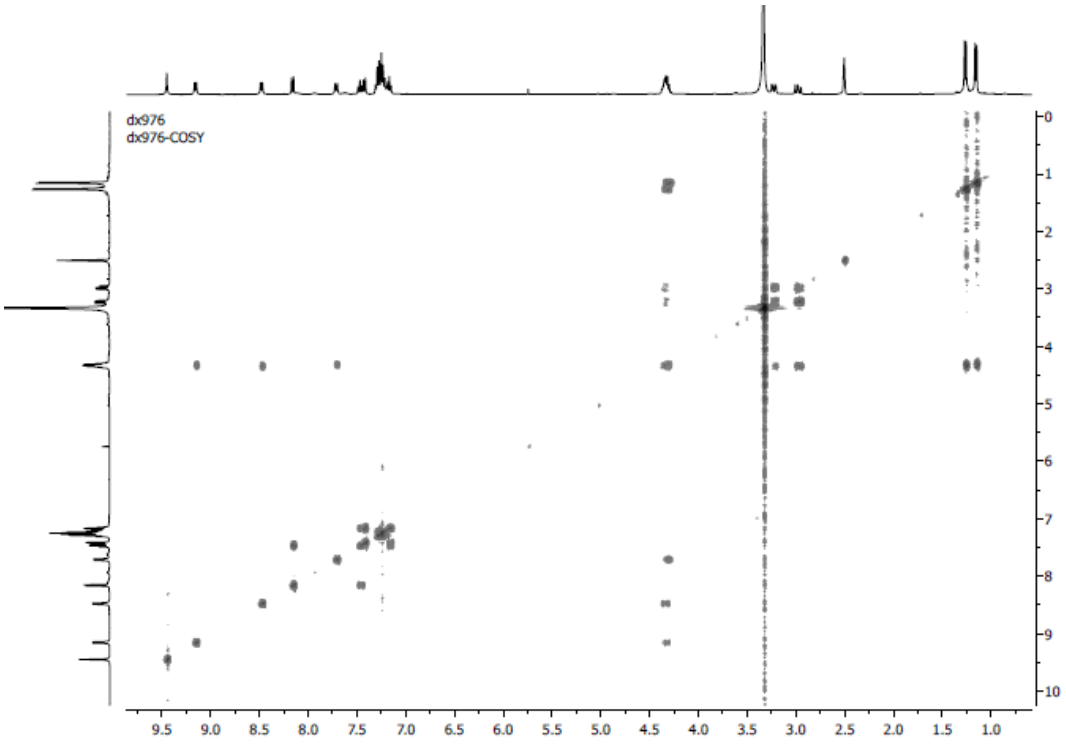
^{13}C NMR



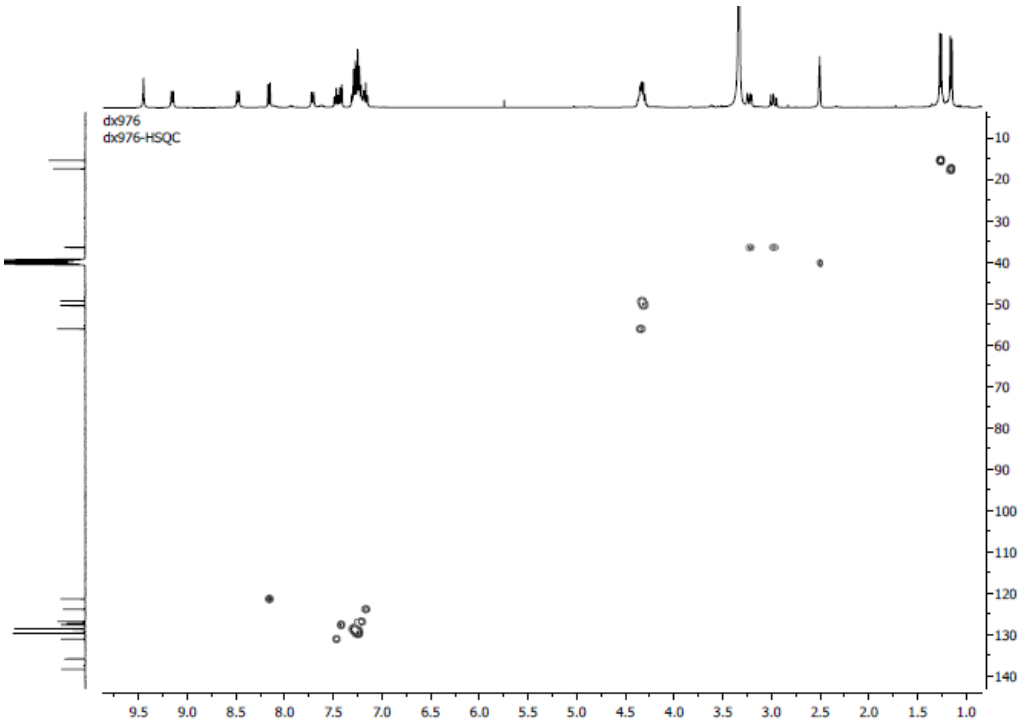
DEPT135


210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0

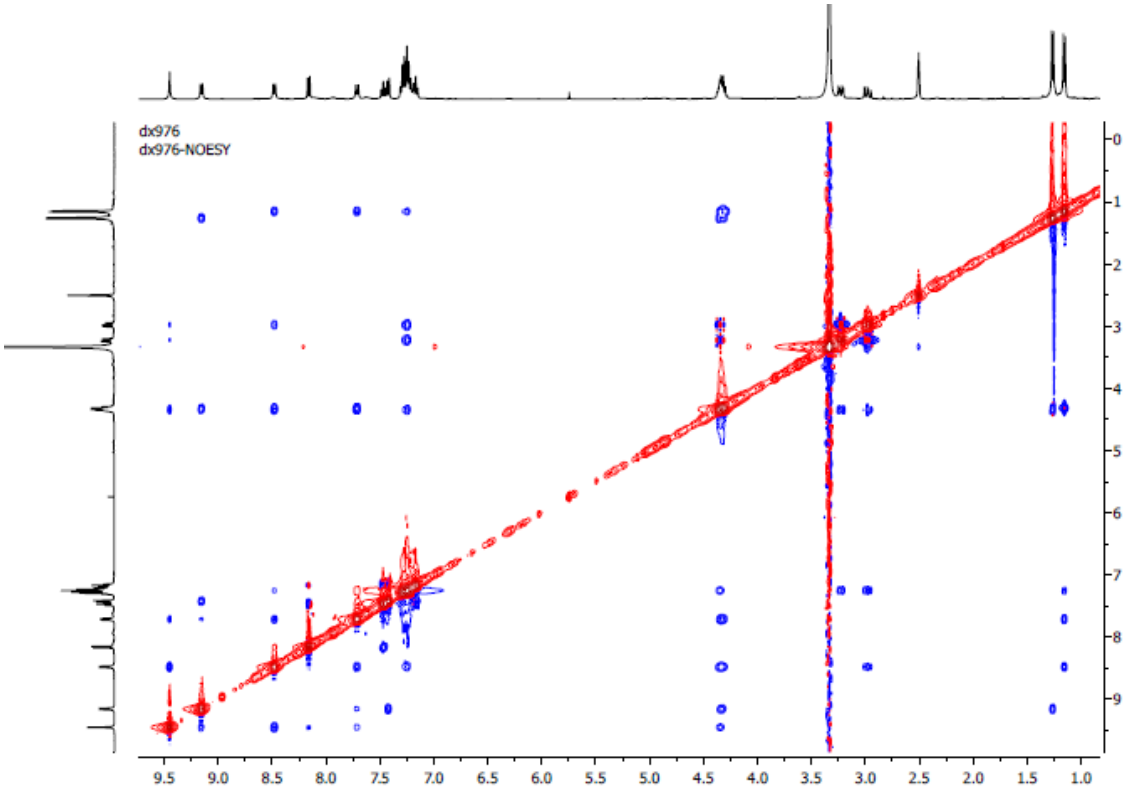
^{1}H-^{1}H COSY



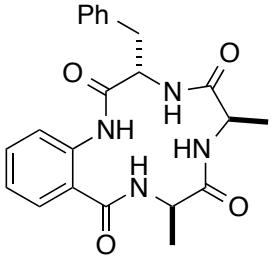
^{1}H-^{13}C HSQC


dx976
dx976-HSQC
10
20
30
40
50
60
70
80
90
100
110
120
130
140
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0

NOESY


dx976
dx976-NOESY
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0
1
2
3
4
5
6
7
8
9

(3S,6R,9R)-3-Benzyl-6,9-dimethyl-3,4,6,7,9,10-hexahydro-1H-benzo[k][1,4,7,10]tetraazacyclotridecine-2,5,8,11-tetraone (DDL-**1aaf**)



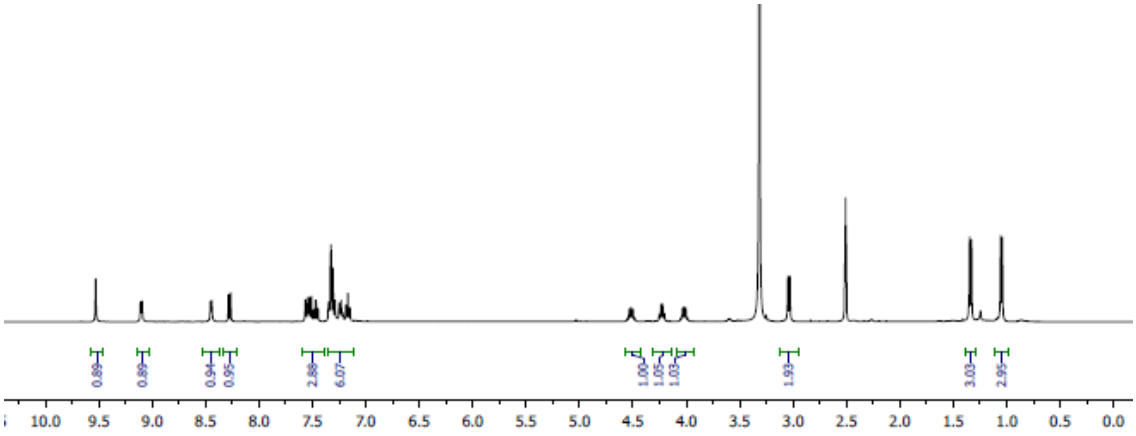
White solid, 44 % over three steps; mp > 270 °C;

^{1}H-NMR (400 MHz, DMSO-d_6) δ 9.53 (s, 1H), 9.10 (d, J = 5.6 Hz, 1H), 8.45 (d, J = 5.2 Hz, 1H), 8.28 (d, J = 8.1 Hz, 1H), 7.60-7.39 (m, 3H), 7.36-7.11 (m, 6H), 4.55-4.48 (m, 1H), 4.26-4.19 (m, 1H), 4.05-3.99 (m, 1H), 3.04 (d, J = 7.4 Hz, 2H), 1.34 (d, J = 7.4 Hz, 3H), 1.05 (d, J = 6.6 Hz, 3H);

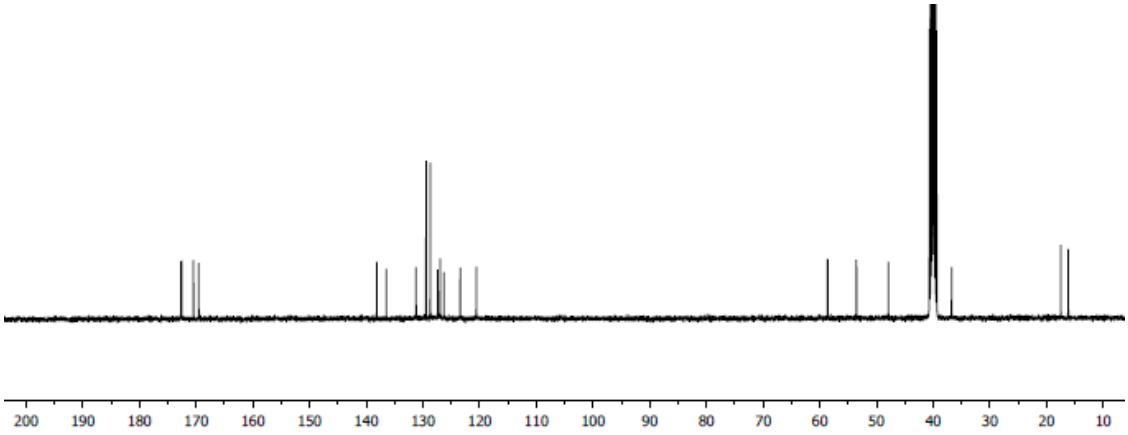
^{13}C-NMR (101 MHz, DMSO-d_6) δ 172.7, 172.5, 170.4, 169.5, 138.1, 136.4, 131.2, 129.4, 128.7, 127.3, 126.9, 126.3, 123.4, 120.5, 58.6, 53.5, 47.9, 36.7, 17.4, 16.1;

HRMS (ESI+) m/z calcd for $C_{22}H_{24}N_4O_4Na$ $(M+Na)^+$ 431.1695; found 431.1716.

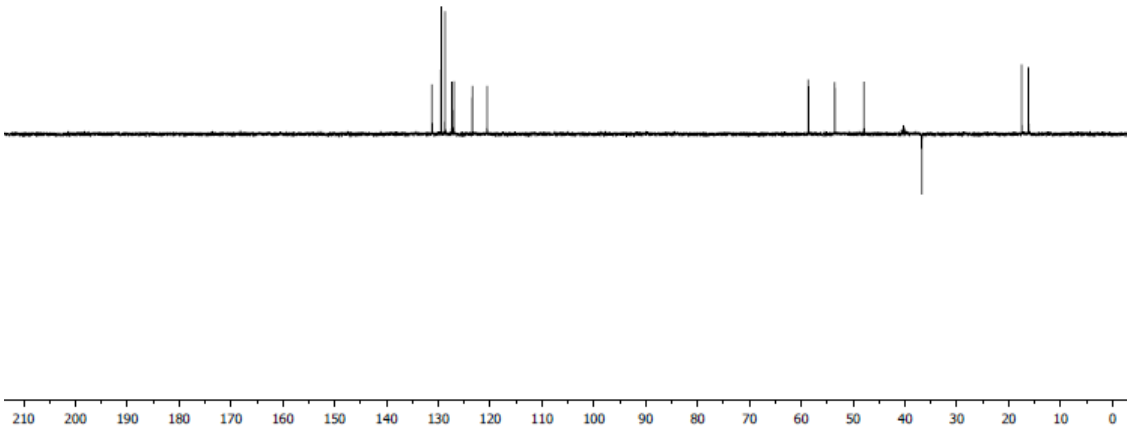
^{1}H NMR



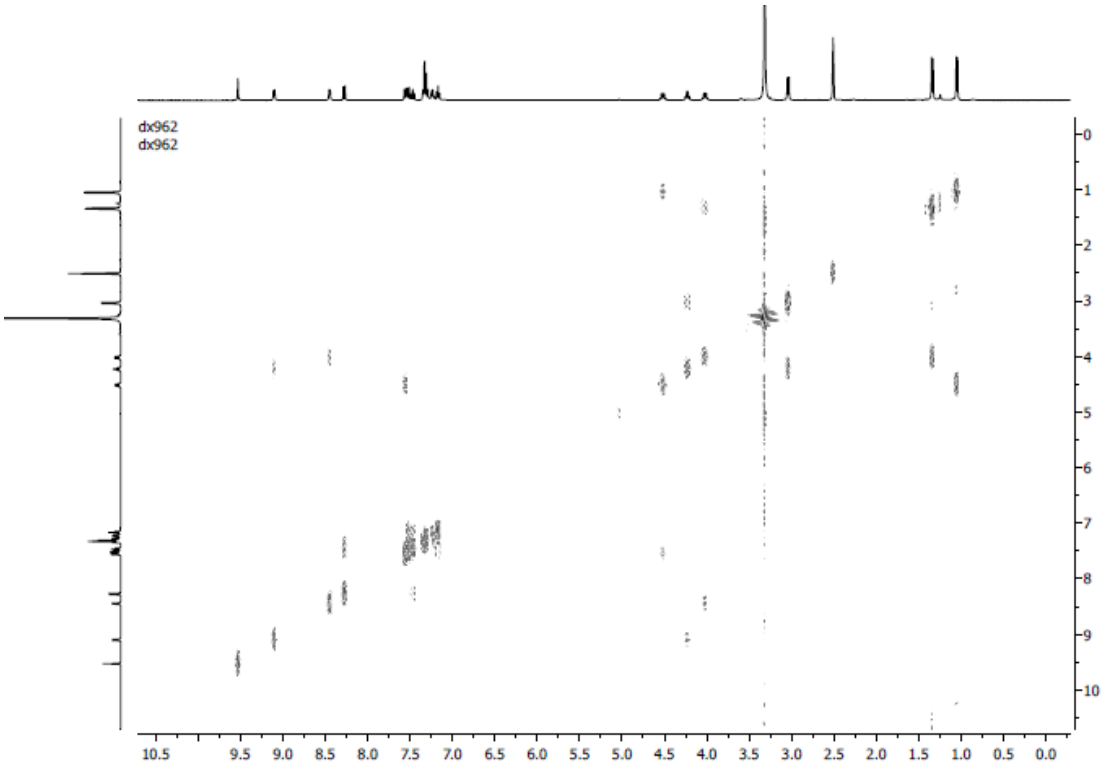
^{13}C NMR



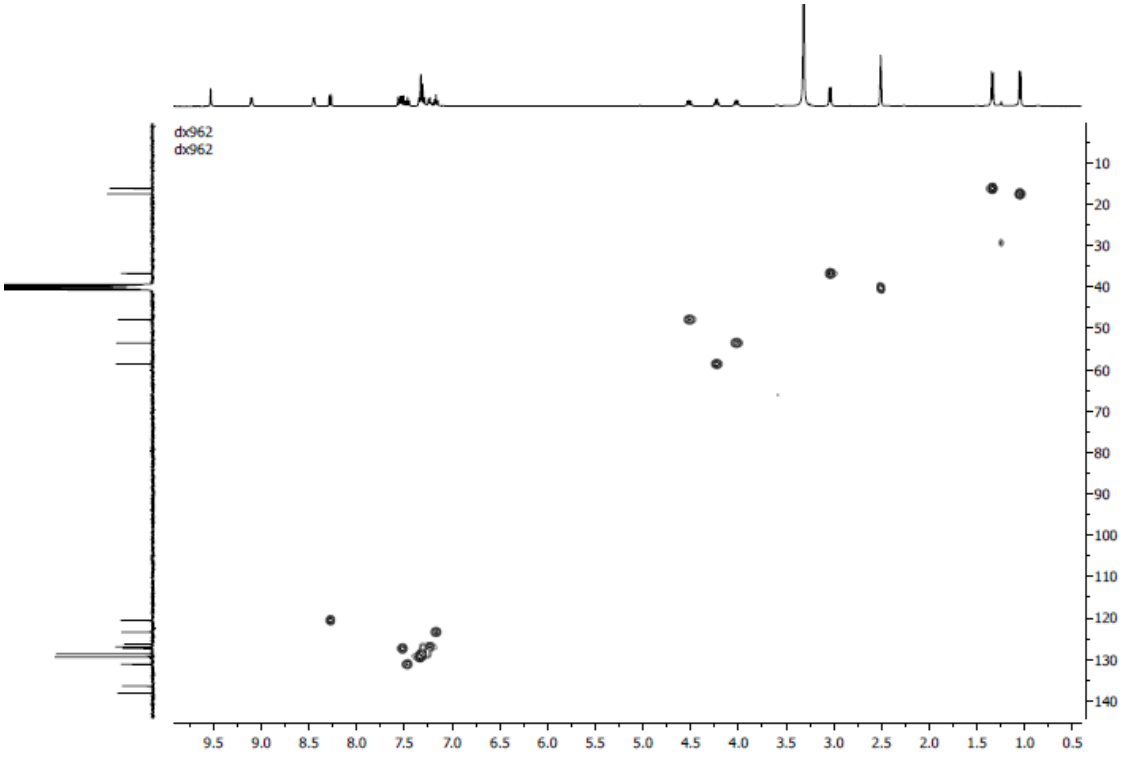
DEPT135


210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0

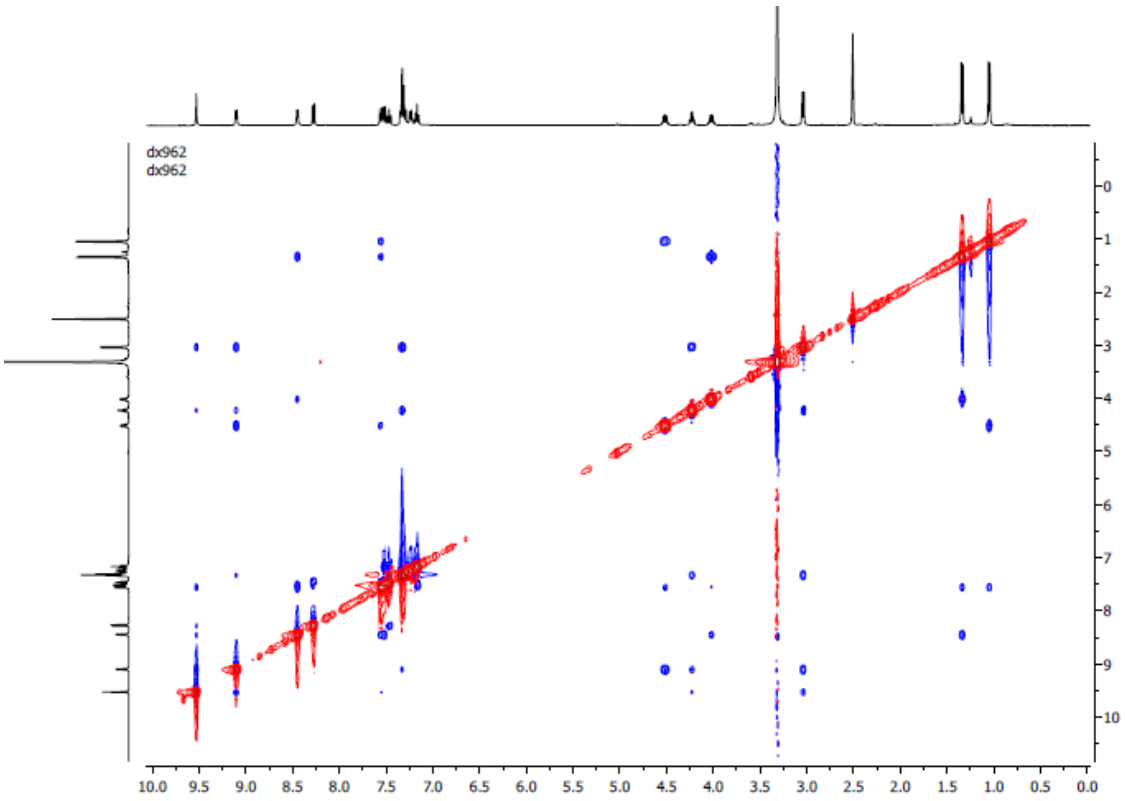
^{1}H-^{1}H COSY



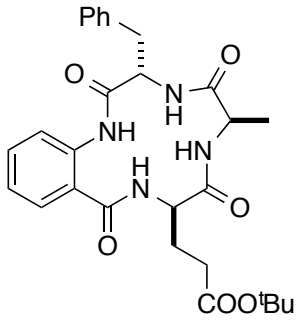
^{1}H-^{13}C HSQC


dx962
dx962
10
20
30
40
50
60
70
80
90
100
110
120
130
140
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5

NOESY


dx962
dx962
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
0
1
2
3
4
5
6
7
8
9
10

tert-Butyl 3-((3S,6R,9R)-3-Benzyl-6-methyl-2,5,8,11-tetraoxo-2,3,4,5,6,7,8,9,10,11-decahydro-1H-benzo[k][1,4,7,10]tetraazacyclotridecin-9-yl)propanoate (DDL-**1e'af**)



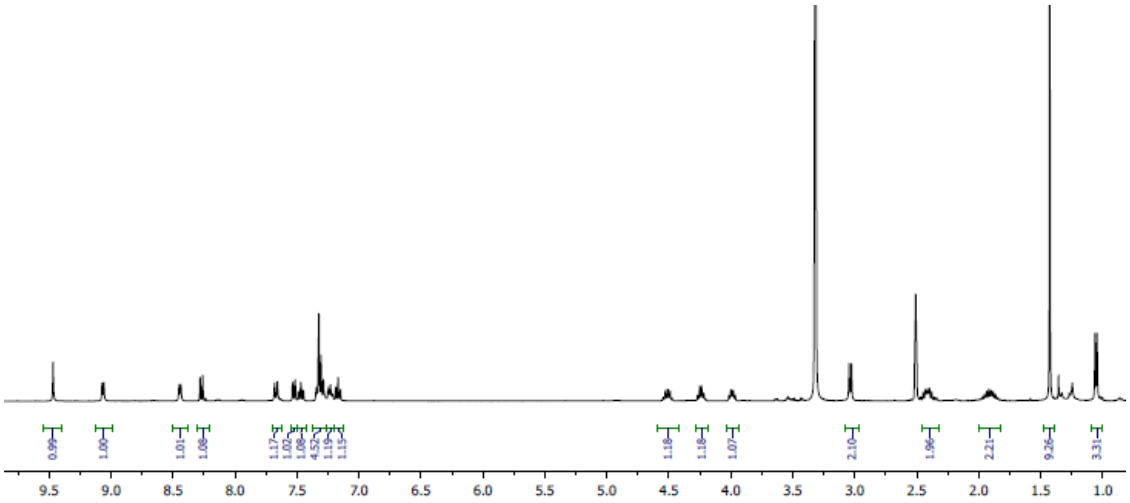
White solid, 35 % over three steps; mp > 270 °C, color change observed at above 230 °C;

^{1}H-NMR (400 MHz, DMSO-d_6) δ 9.47 (s, 1H), 9.06 (d, J = 5.8 Hz, 1H), 8.44 (d, J = 5.6 Hz, 1H), 8.27 (d, J = 8.1 Hz, 1H), 7.67 (d, J = 8.8 Hz, 1H), 7.53 (dd, J = 7.6, 1.2 Hz, 1H), 7.47 (t, J = 7.8 Hz, 1H), 7.35-7.24 (m, 4H), 7.24-7.19 (m, 1H), 7.17 (t, J = 7.5 Hz, 1H), 4.60-4.41 (m, 1H), 4.29-4.19 (m, 1H), 4.04-3.95 (m, 1H), 3.04 (d, J = 7.4 Hz, 2H), 2.46-2.32 (m, 2H), 1.99-1.83 (m, 2H), 1.43 (s, 9H), 1.05 (d, J = 6.7 Hz, 3H);

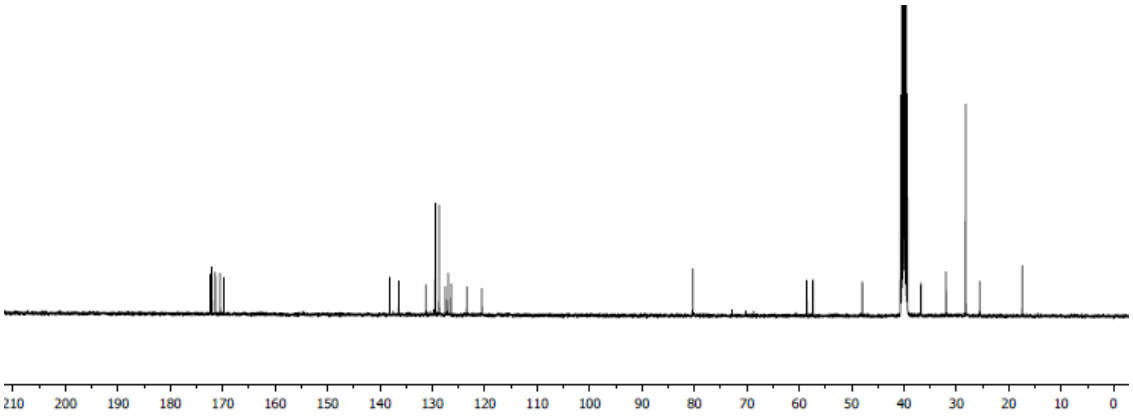
^{13}C-NMR (101 MHz, DMSO-d_6) δ 172.4, 172.0, 171.4, 170.4, 169.7, 138.1, 136.4, 131.2, 129.4, 128.7, 127.5, 126.9, 126.4, 123.4, 120.5, 80.3, 58.5, 57.3, 47.9, 36.7, 31.9, 28.2, 25.5, 17.4;

HRMS (ESI-) m/z calcd for $C_{28}H_{33}N_4O_6$ $(M-H)^-$ 521.2400; found 521.2422.

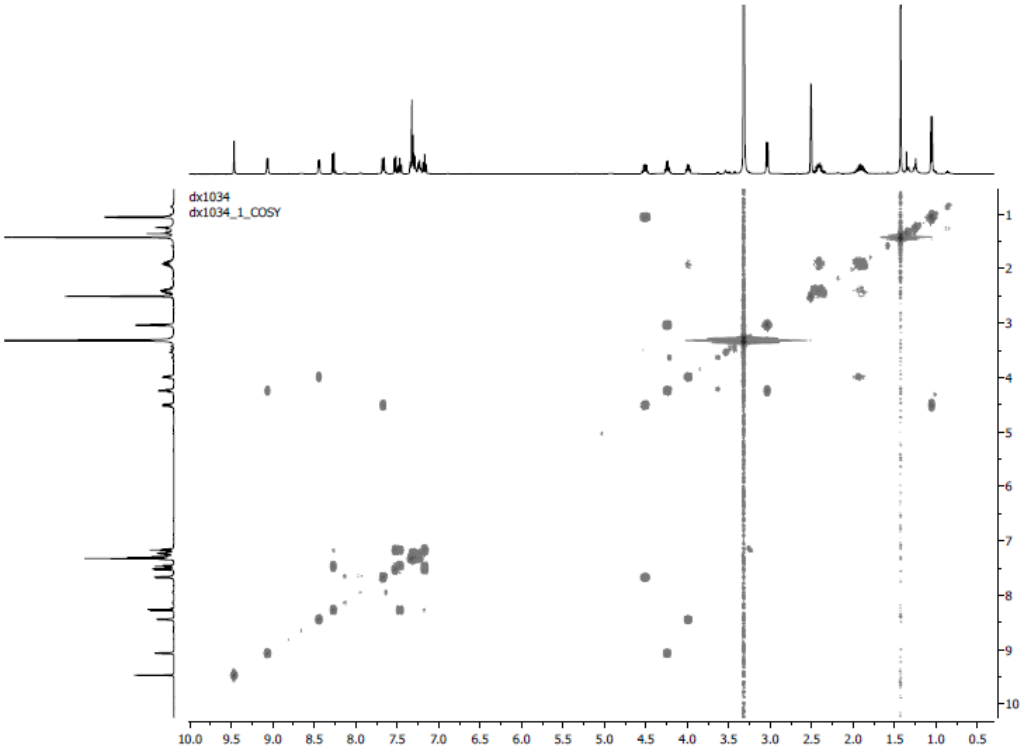
^{1}H NMR



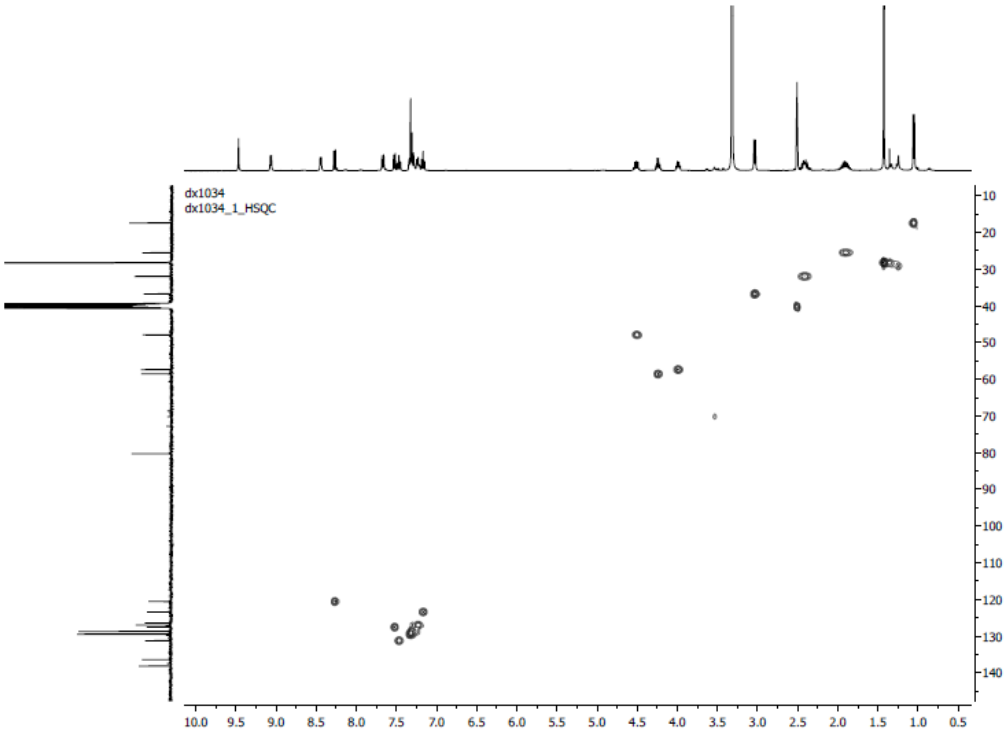
^{13}C NMR



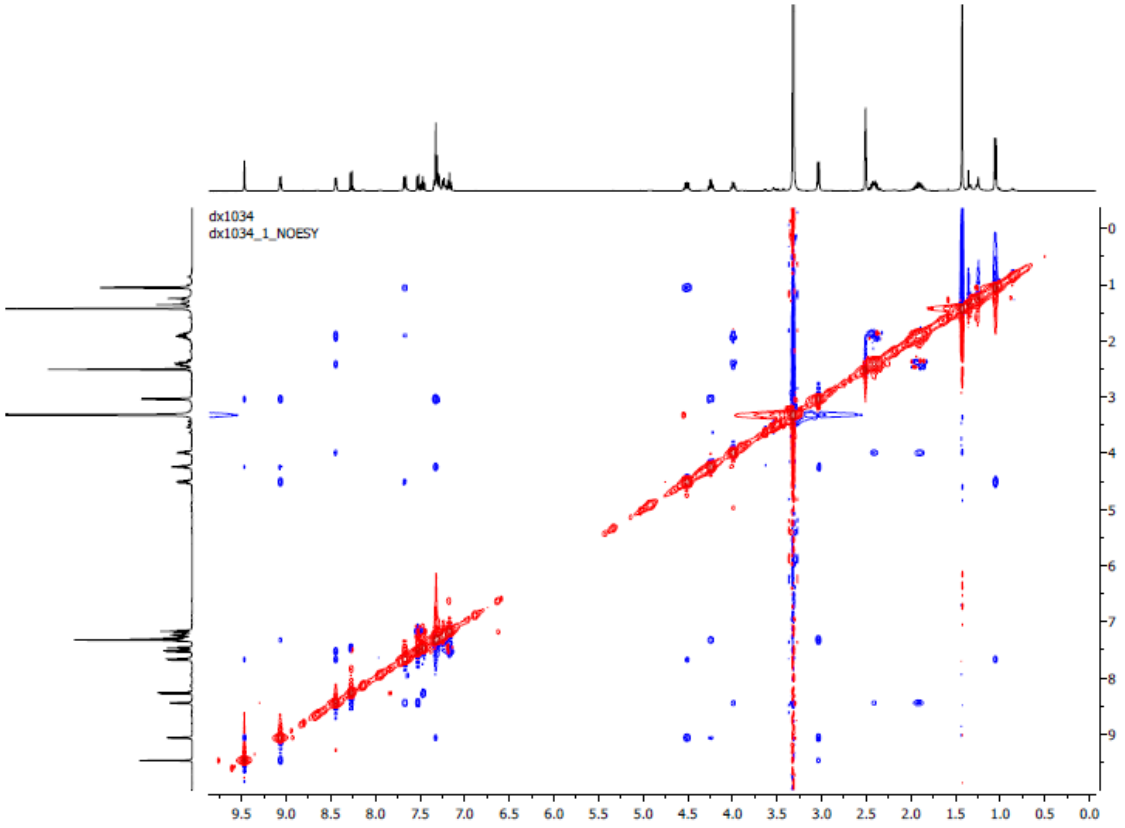
^{1}H-^{1}H COSY



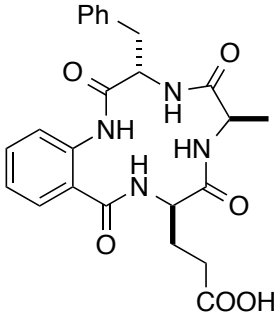
^{1}H-^{13}C HSQC


dx1034
dx1034_1_HSQC
10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5
10 20 30 40 50 60 70 80 90 100 110 120 130 140

NOESY


dx1034
dx1034_1_NOESY
9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
0 1 2 3 4 5 6 7 8 9

3-((3S,6R,9R)-3-Benzyl-6-methyl-2,5,8,11-tetraoxo-2,3,4,5,6,7,8,9,10,11-decahydro-1H-benzo[k][1,4,7,10]tetraazacyclotridecin-9-yl)propanoic acid (DDL-**1eaf**)



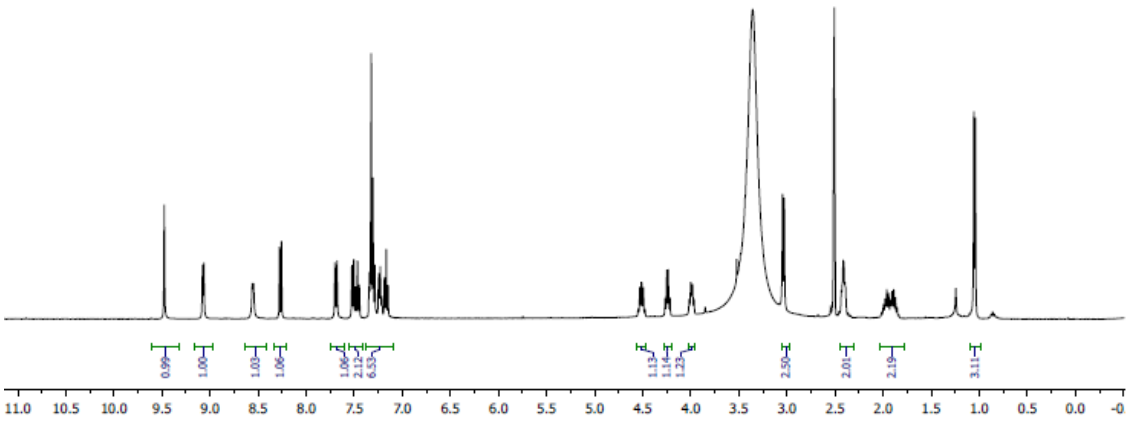
White solid, 87 %; mp > 270 °C;

^{1}H-NMR (400 MHz, DMSO-d_6) δ 9.48 (s, 1H), 9.07 (d, J = 5.7 Hz, 1H), 8.55 (d, J = 5.1 Hz, 1H), 8.27 (d, J = 8.2 Hz, 1H), 7.69 (d, J = 8.8 Hz, 1H), 7.56-7.41 (m, 2H), 7.37-7.10 (m, 7H), 4.58-4.48 (m, 1H), 4.29-4.22 (m, 1H), 4.05-3.96 (m, 1H), 3.04 (d, J = 7.4 Hz, 2H), 2.46-2.38 (m, 2H), 2.03-1.78 (m, 2H), 1.05 (d, J = 6.6 Hz, 3H);

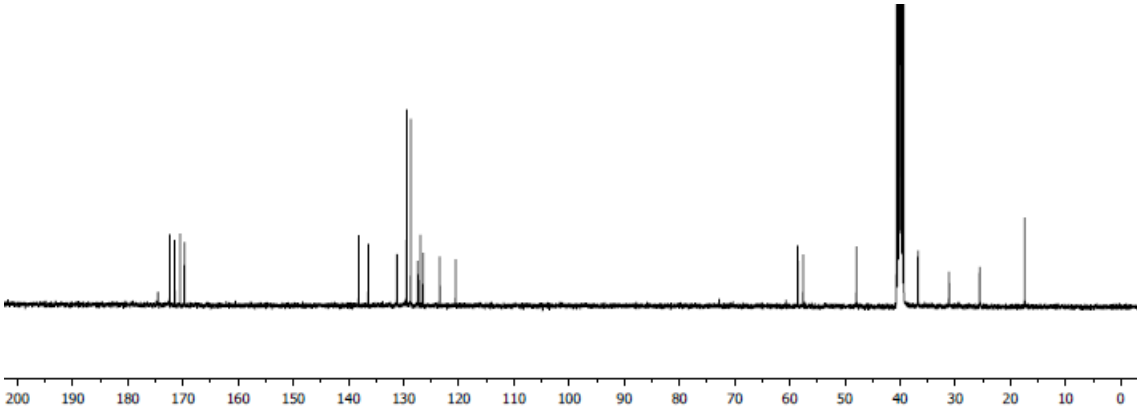
^{13}C-NMR (101 MHz, DMSO-d_6) δ 174.5, 172.4, 171.5, 170.5, 169.7, 138.1, 136.4, 131.2, 129.4, 128.7, 127.4, 126.9, 126.4, 123.4, 120.5, 58.5, 57.5, 47.9, 36.7, 31.1, 25.5, 17.4;

HRMS (ESI-) m/z calcd for $C_{24}H_{25}N_4O_6$ $(M-H)^-$ 465.1774; found 465.1786.

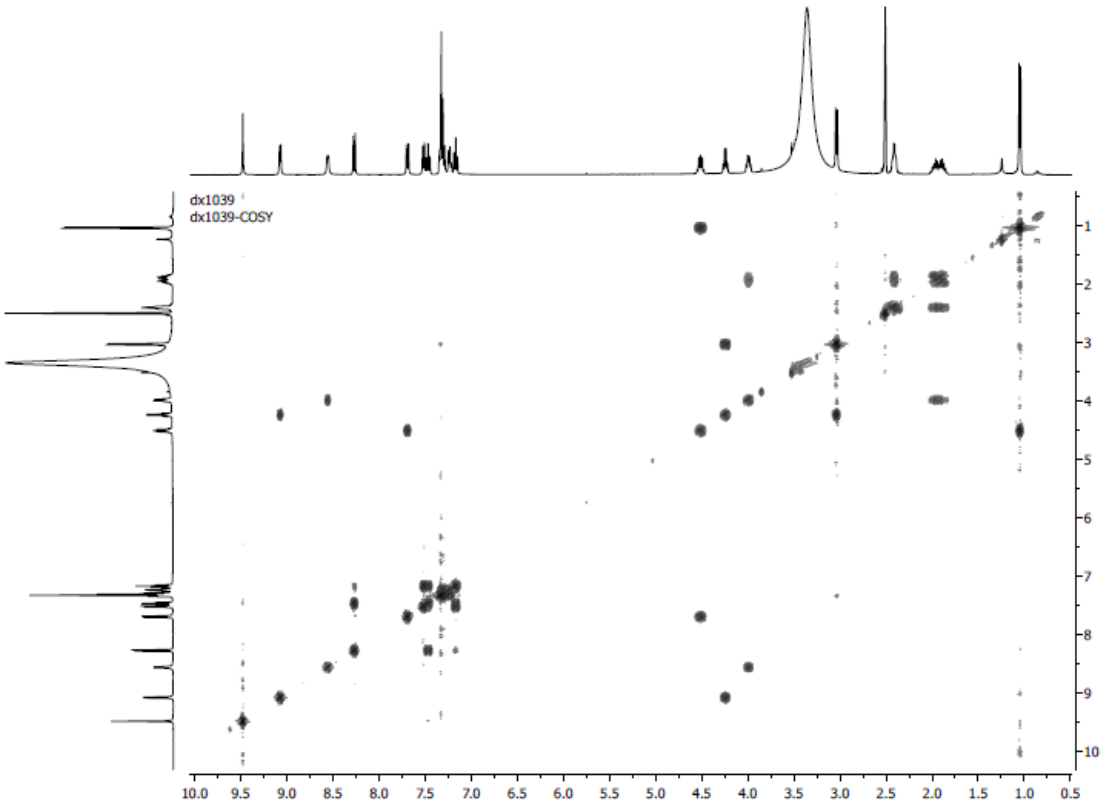
^{1}H NMR


0.99
1.00
1.03
1.06
1.06
2.12
6.53
1.13
1.14
1.23
2.50
2.01
2.19
3.11
11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0

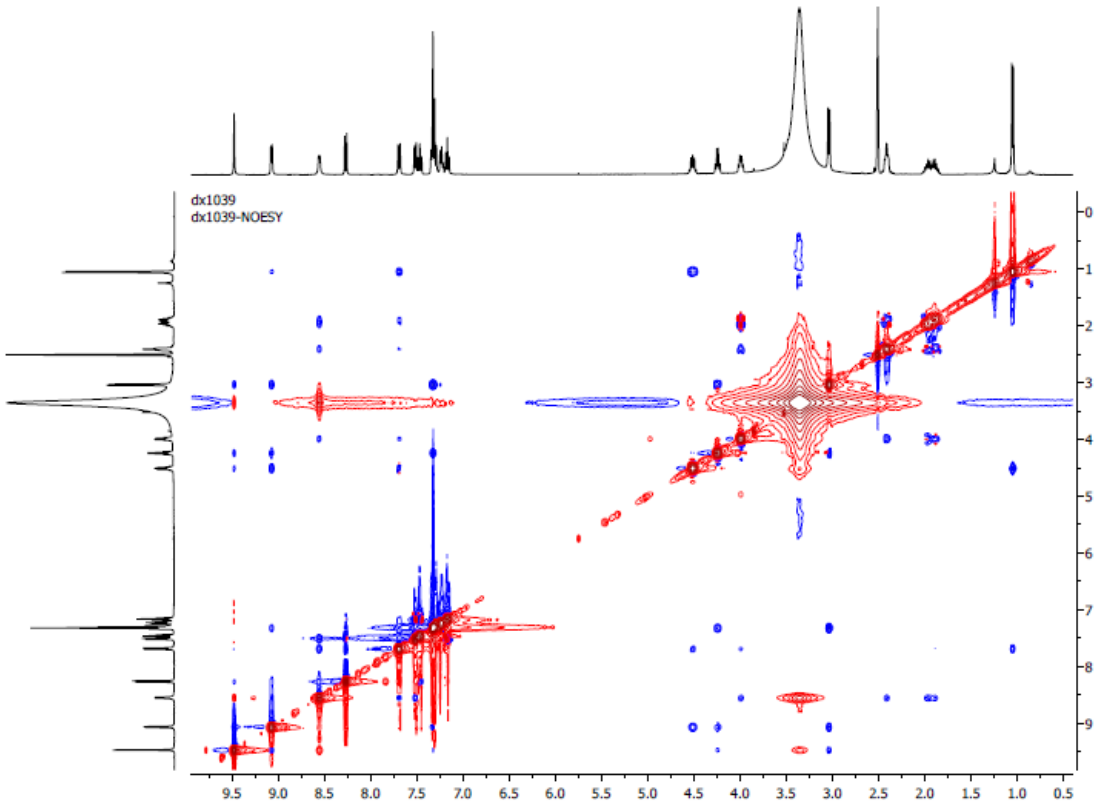
^{13}C NMR


200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0

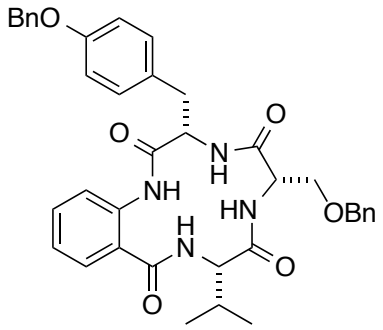
^{1}H-^{1}H COSY


dx1039
dx1039-COSY
10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5
1 2 3 4 5 6 7 8 9 10

NOESY


dx1039
dx1039-NOESY
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0
1
2
3
4
5
6
7
8
9

(3S,6S,9S)-3-(4-(Benzyloxy)benzyl)-6-((benzyloxy)methyl)-9-isopropyl-3,4,6,7,9,10-hexahydro-1H-benzo[k][1,4,7,10]tetraazacyclotridecine-2,5,8,11-tetraone (LLL-**1vs'y'**)



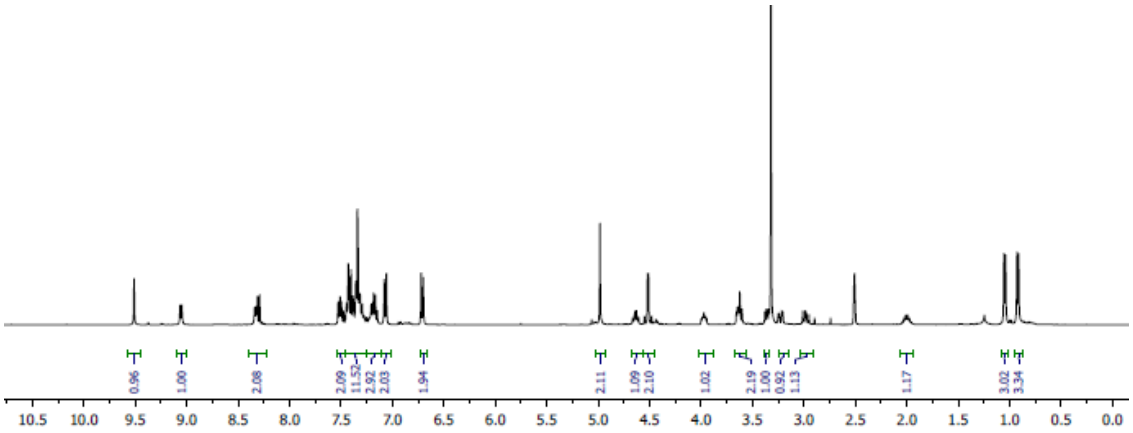
White solid, 45 % over three steps; mp > 270 °C;

^{1}H-NMR (400 MHz, DMSO-d_6) δ 9.51 (s, 1H), 9.06 (d, J = 6.8 Hz, 1H), 8.36-8.28 (m, 2H), 7.54-7.47 (m, 2H), 7.47-7.25 (m, 10H), 7.25-7.11 (m, 2H), 7.07 (d, J = 8.5 Hz, 2H), 6.71 (d, J = 8.5 Hz, 2H), 4.98 (s, 2H), 4.69-4.60 (m, 1H), 4.57-4.46 (m, 2H), 4.03-3.88 (m, 1H), 3.67-3.58 (m, 2H), 3.36 (dd, J = 9.2, 5.8 Hz, 1H), 3.24-3.15 (m, 1H), 2.98 (dd, J = 14.1, 10.1 Hz, 1H), 2.05-1.97 (m, 1H), 1.05 (d, J = 6.6 Hz, 3H), 0.92 (d, J = 6.7 Hz, 3H);

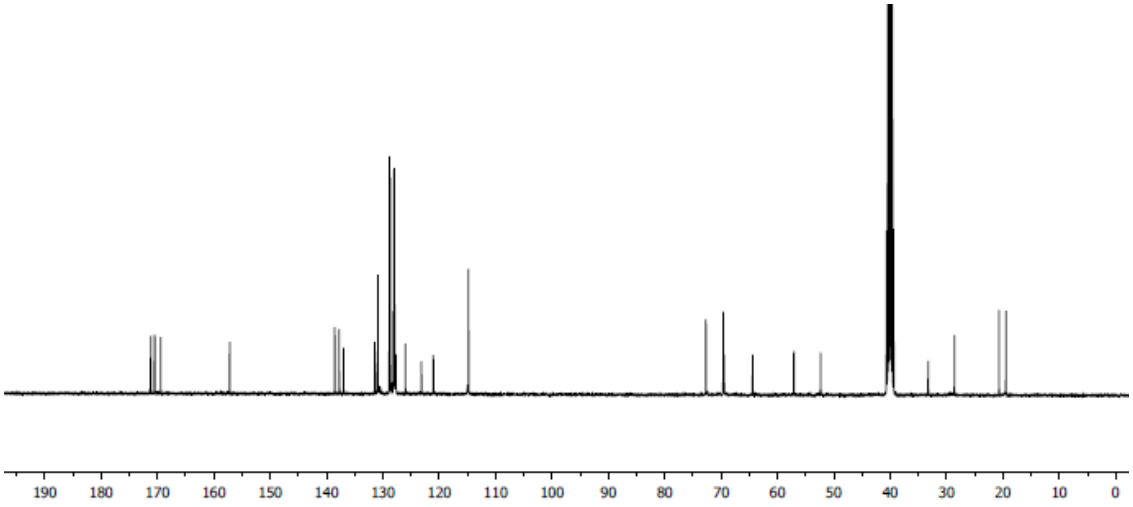
^{13}C-NMR (101 MHz, DMSO-d_6) δ 171.2, 170.5, 170.4, 169.4, 157.1, 138.5, 137.7, 136.9, 131.4, 131.3, 130.9, 128.8, 128.6, 128.2, 128.0, 127.9, 127.8, 127.7, 126.0, 123.1, 121.0, 114.8, 72.7, 69.6, 69.4, 64.4, 57.0, 52.3, 33.2, 28.6, 20.6, 19.4;

HRMS (ESI-) m/z calcd for $C_{38}H_{39}N_4O_6$ $(M-H)^-$ 647.2870; found 647.2850.

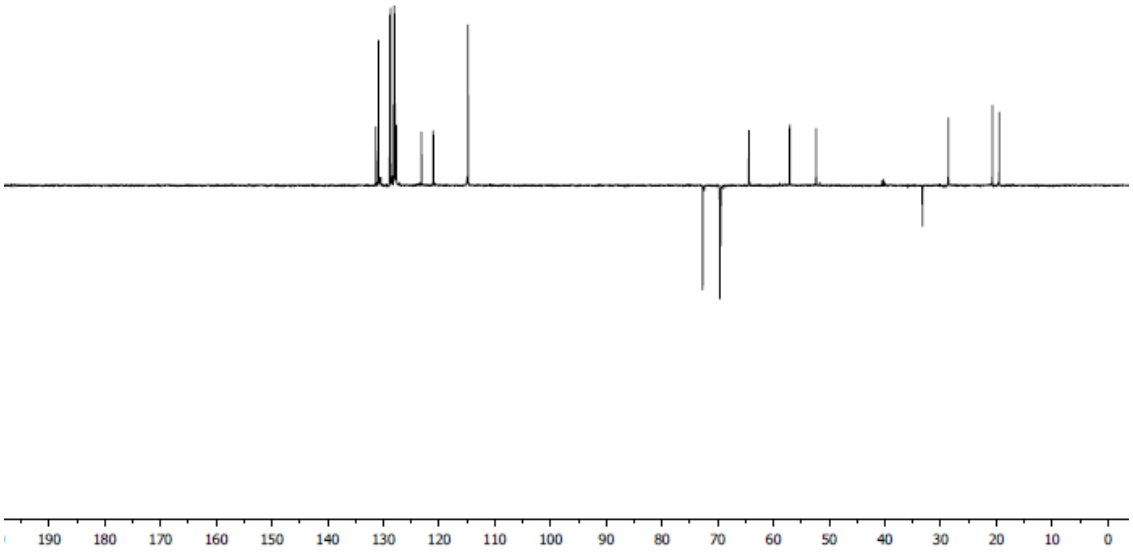
^{1}H NMR



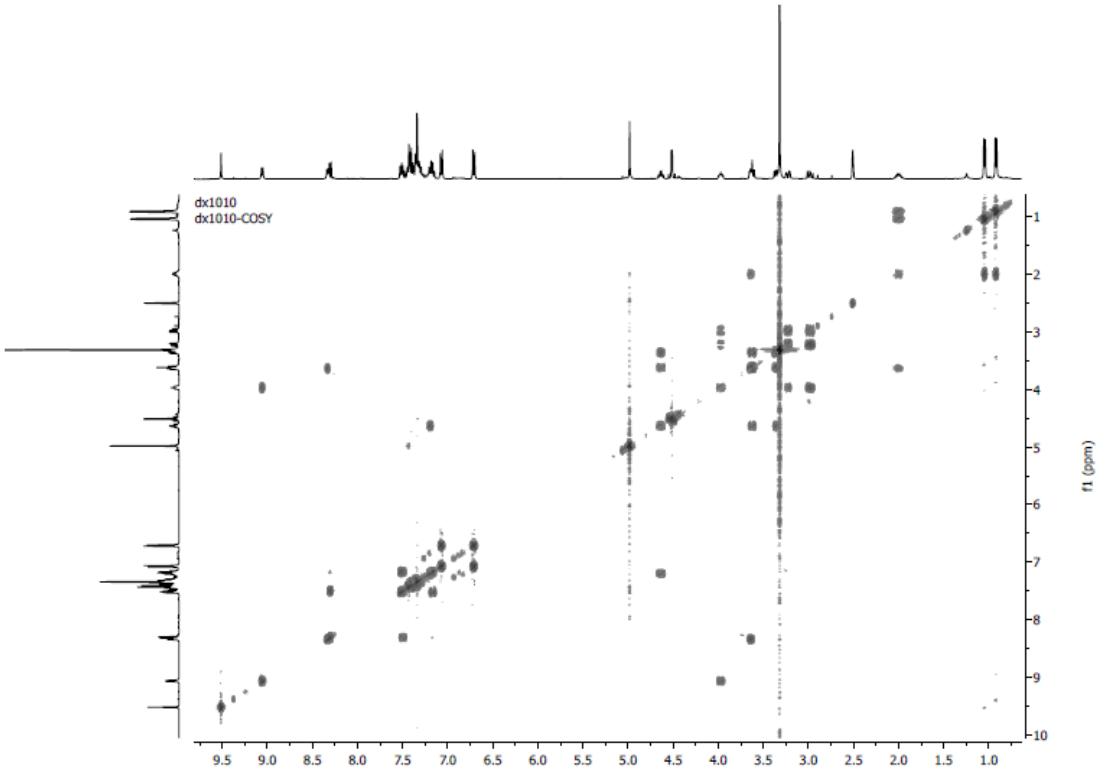
^{13}C NMR



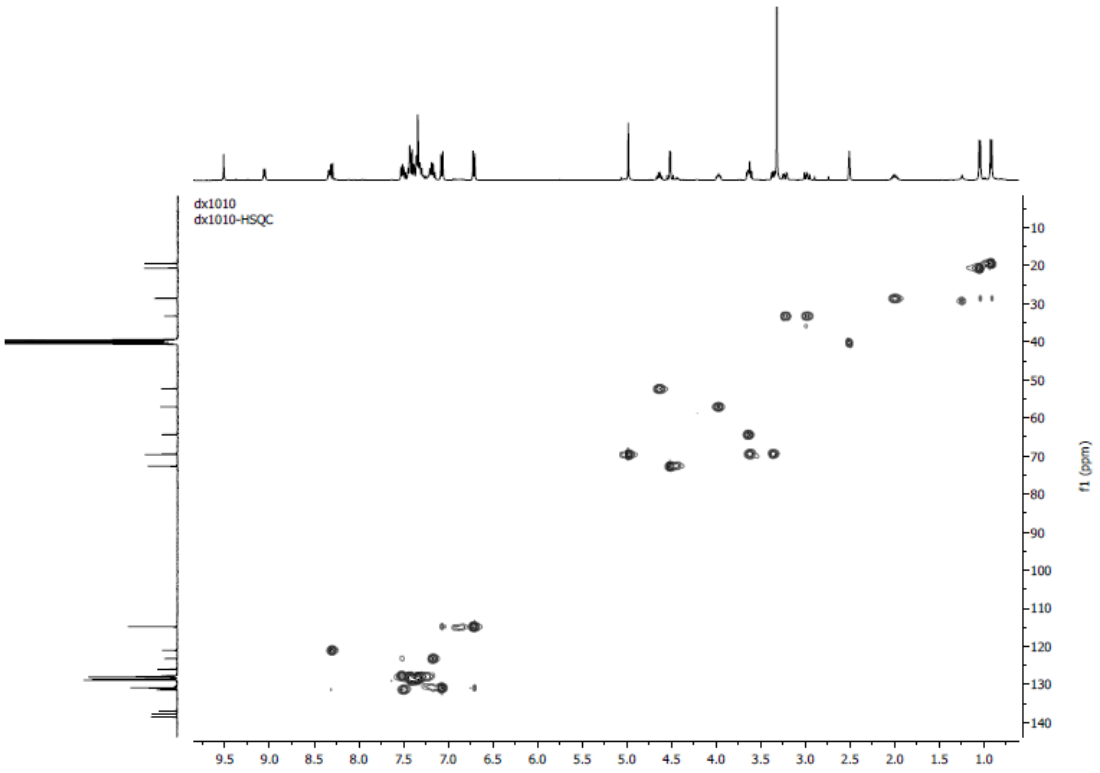
DEPT135


190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0

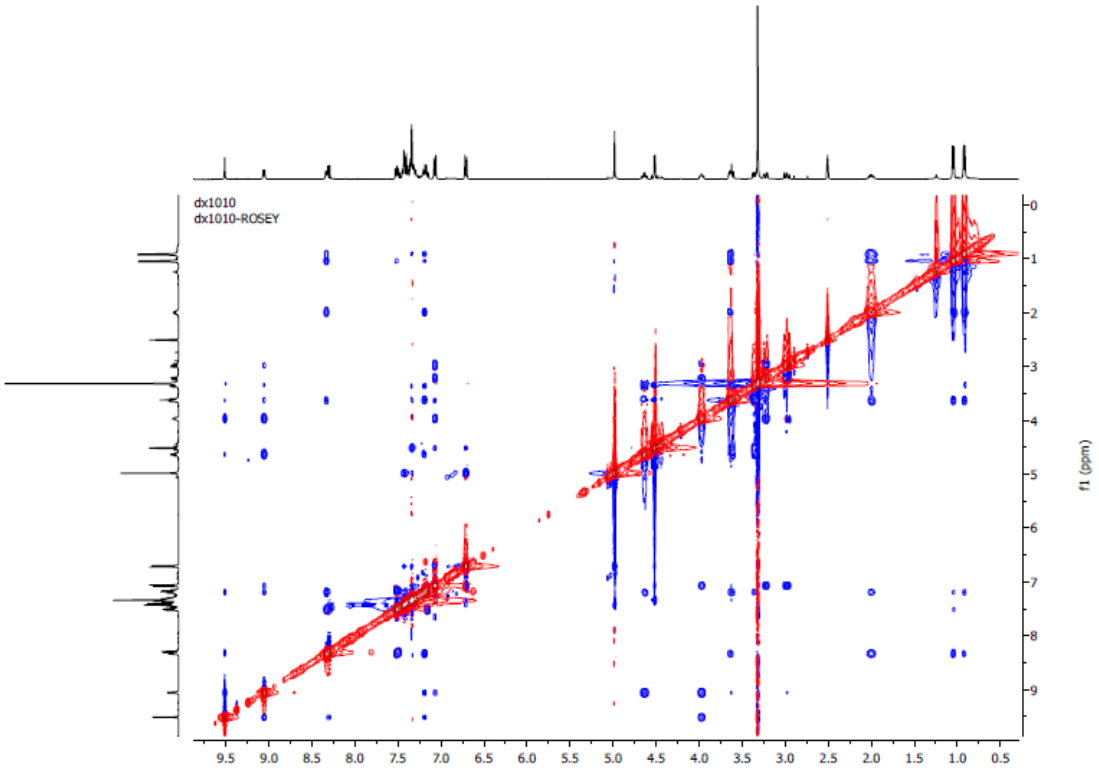
^{1}H-^{1}H COSY


dx1010
dx1010-COSY
9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0
1 2 3 4 5 6 7 8 9 10
f1 (ppm)

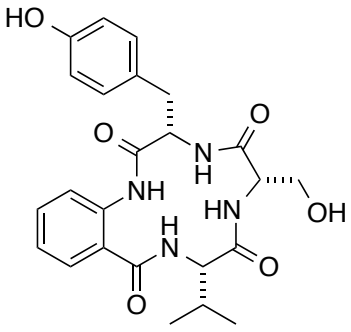
^{1}H-^{13}C HSQC



ROESY



(3S,6S,9S)-3-(4-Hydroxybenzyl)-6-(hydroxymethyl)-9-isopropyl-3,4,6,7,9,10-hexahydro-1H-benzo[k][1,4,7,10]tetraazacyclotridecine-2,5,8,11-tetraone (LLL-**1vsy**)



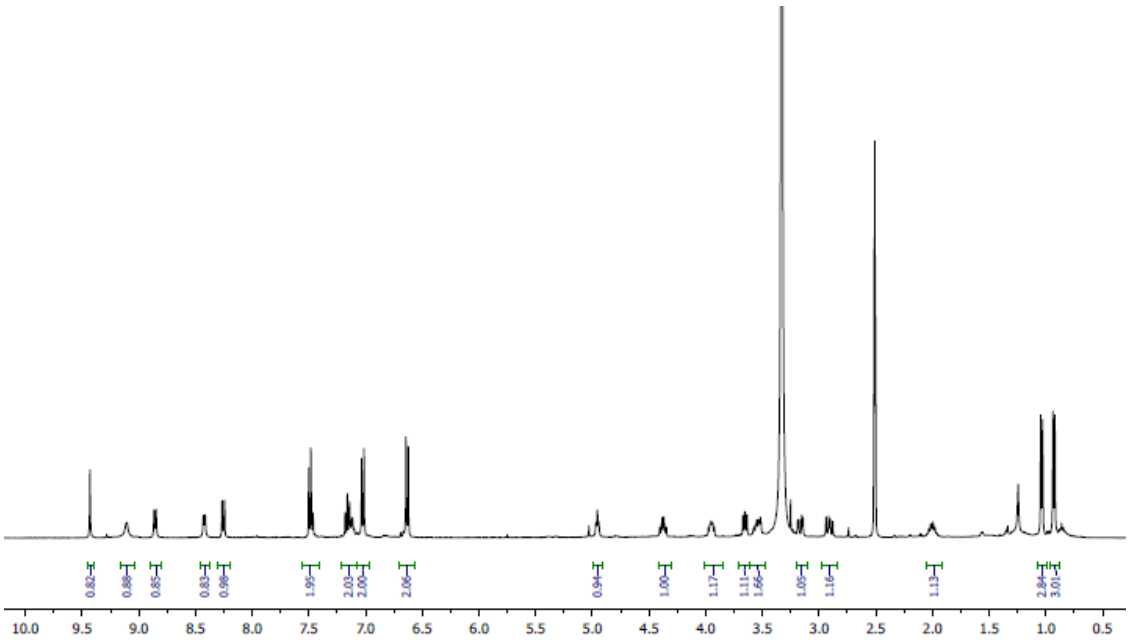
White solid, 87 %; mp > 270 °C, color change observed at above 193 °C;

^{1}H-NMR (400 MHz, DMSO-d_6) δ 9.43 (s, 1H), 9.11 (s, 1H), 8.86 (d, J = 7.1 Hz, 1H), 8.42 (d, J = 6.0 Hz, 1H), 8.25 (d, J = 8.4 Hz, 1H), 7.53-7.45 (m, 2H), 7.21-7.08 (m, 2H), 7.02 (d, J = 8.4 Hz, 2H), 6.63 (d, J = 8.5 Hz, 2H), 4.96 (t, J = 5.3 Hz, 1H), 4.41-4.33 (m, 1H), 3.98-3.91 (m, 1H), 3.65 (dd, J = 8.8, 6.3 Hz, 1H), 3.61-3.47 (m, 2H), 3.17 (dd, J = 14.0, 4.4 Hz, 1H), 2.91 (dd, J = 14.0, 9.5 Hz, 1H), 2.05-1.92 (m, 1H), 1.04 (d, J = 6.6 Hz, 3H), 0.93 (d, J = 6.7 Hz, 3H);

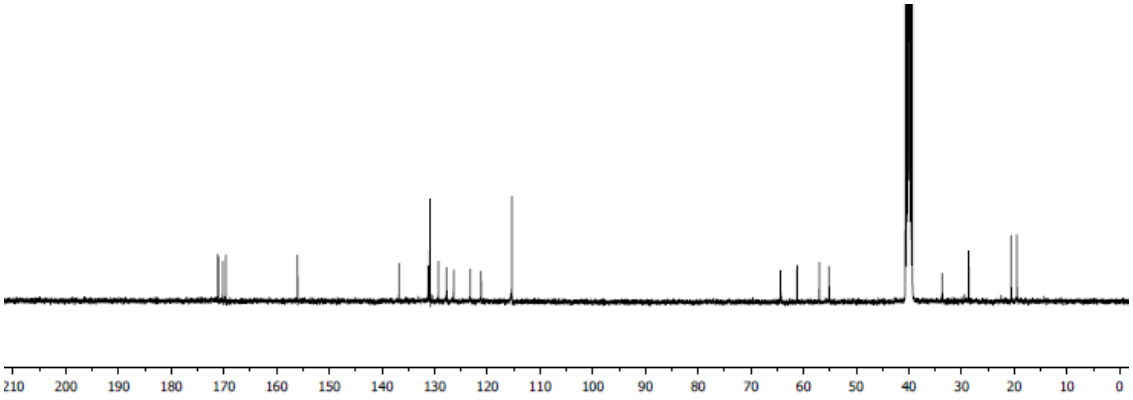
^{13}C-NMR (101 MHz, DMSO-d_6) δ 171.2, 171.1, 170.2, 169.6, 156.0, 136.7, 131.2, 130.9, 129.3, 127.7, 126.4, 123.2, 121.2, 115.3, 64.4, 61.2, 57.0, 55.2, 33.7, 28.6, 20.6, 19.5;

HRMS (ESI-) m/z calcd for $C_{24}H_{27}N_4O_6$ $(M-H)^-$ 467.1931; found 467.1919.

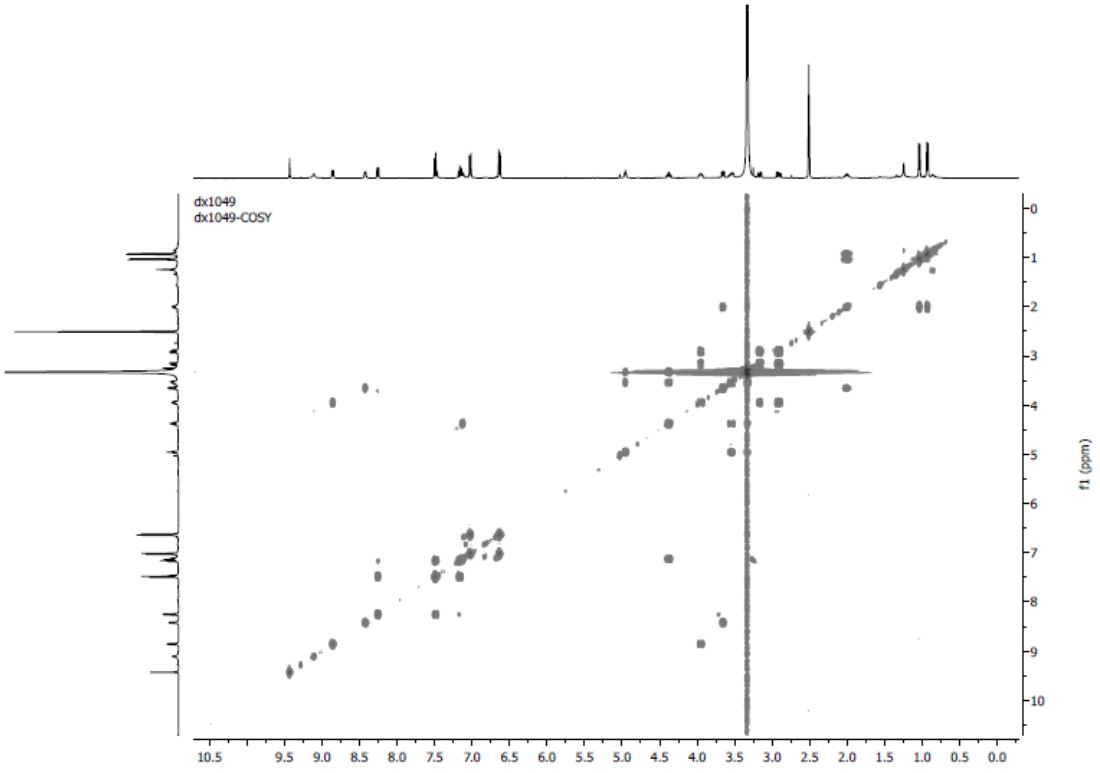
^{1}H NMR



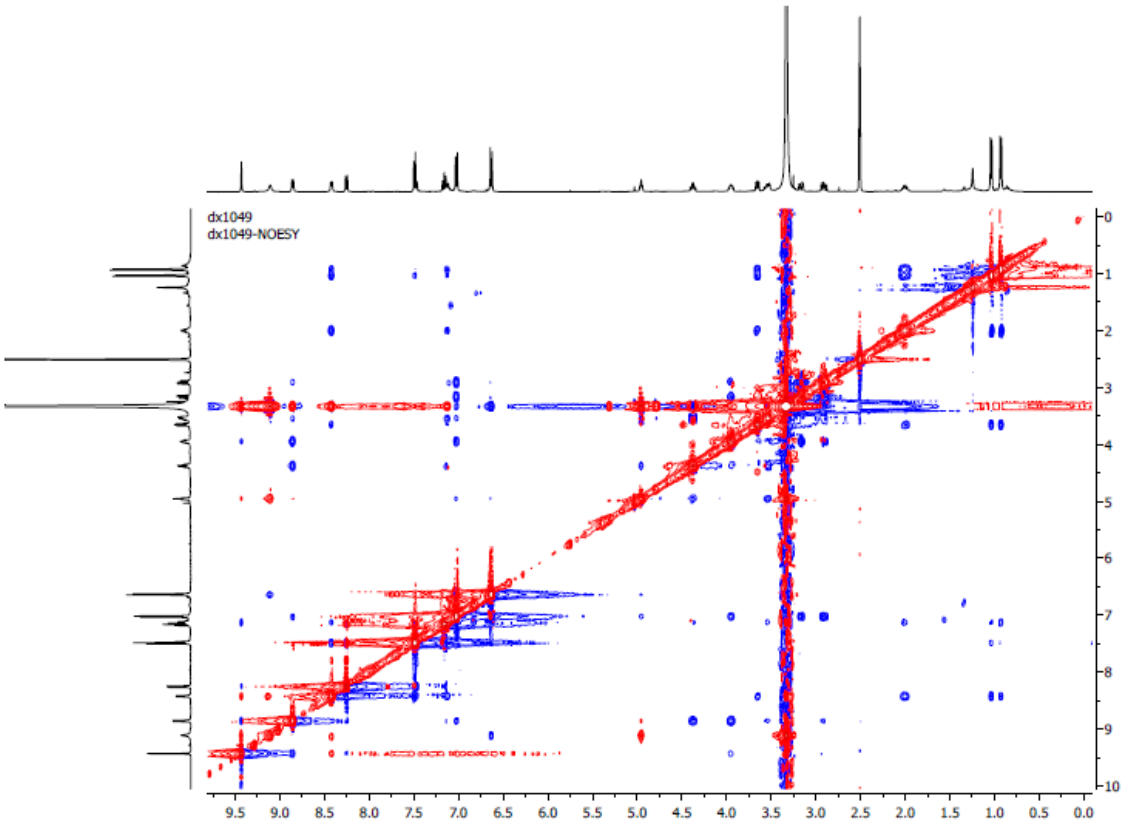
^{13}C NMR


210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0

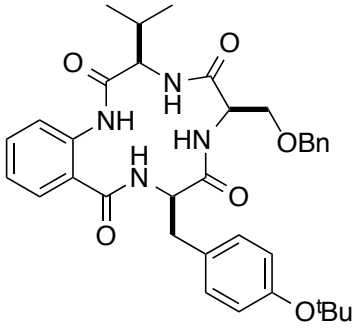
^{1}H-^{1}H COSY


dx1049
dx1049-COSY
f1 (ppm)
10.5
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0

NOESY


dx1049
dx1049-NOESY
9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
0 1 2 3 4 5 6 7 8 9 10

(3R,6R,9R)-6-((Benzyloxy)methyl)-9-(4-(tert-butoxy)benzyl)-3-isopropyl-3,4,6,7,9,10-hexahydro-1H-benzo[k][1,4,7,10]tetraazacyclotridecine-2,5,8,11-tetraone (DDD-**1y's'v**)



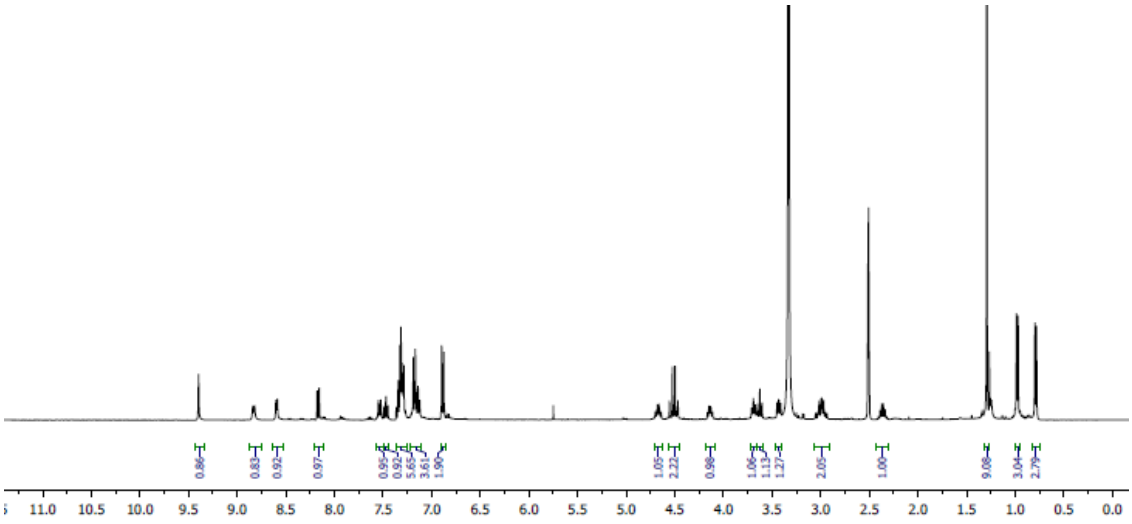
White solid, 31 % over three steps; mp > 270 °C;

^{1}H-NMR (400 MHz, DMSO-d_6) δ 9.40 (s, 1H), 8.83 (d, J = 7.3 Hz, 1H), 8.59 (d, J = 5.4 Hz, 1H), 8.17 (d, J = 8.1 Hz, 1H), 7.54 (d, J = 9.5 Hz, 1H), 7.50-7.44 (m, 1H), 7.36-7.26 (m, 6H), 7.21-7.11 (m, 3H), 6.89 (d, J = 8.4 Hz, 2H), 4.67 (td, J = 9.2, 6.0 Hz, 1H), 4.54 (d, J = 12.0 Hz, 1H), 4.48 (d, J = 12.0 Hz, 1H), 4.18-4.09 (m, 1H), 3.69 (t, J = 7.2 Hz, 1H), 3.63 (t, J = 9.3 Hz, 1H), 3.43 (dd, J = 9.4, 6.0 Hz, 1H), 3.06-2.93 (m, 2H), 2.68-2.59 (m, 1H), 1.29 (s, 9H), 0.98 (d, J = 6.7 Hz, 3H), 0.79 (d, J = 6.8 Hz, 3H);

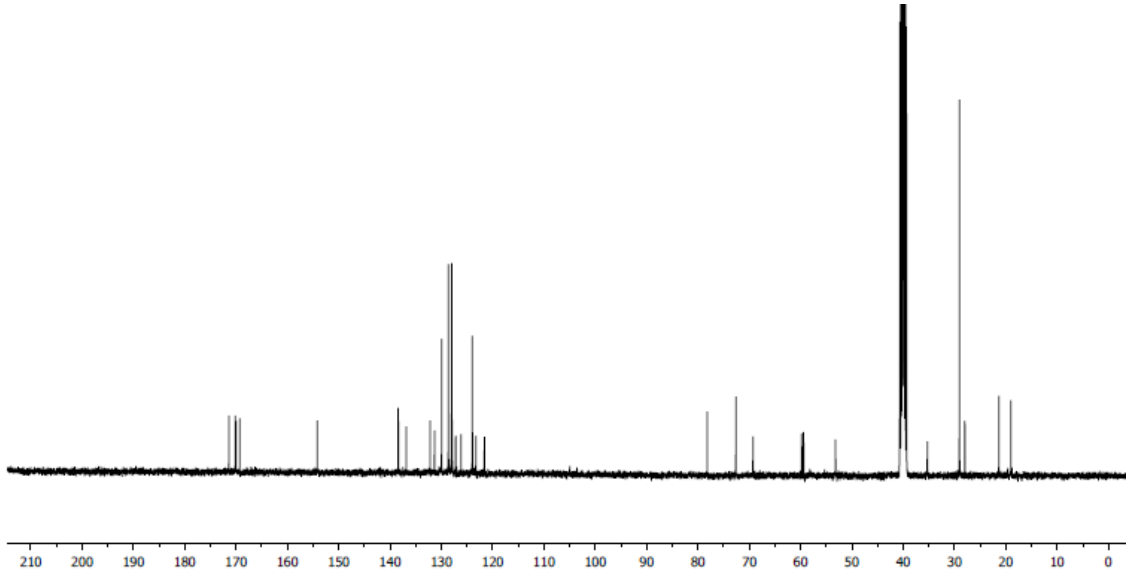
^{13}C-NMR (101 MHz, DMSO-d_6) δ 171.3, 170.1, 170.0, 169.2, 154.1, 138.3, 136.8, 132.1, 131.3, 129.9, 128.6, 128.0, 127.9, 127.2, 126.1, 123.9, 123.2, 121.6, 78.1, 72.6, 69.2, 59.7, 59.4, 53.1, 35.3, 29.0, 27.9, 21.3, 19.0;

HRMS (ESI+) m/z calcd for $C_{35}H_{42}N_4O_6Na$ $(M+Na)^+$ 637.3002; found 637.2988.

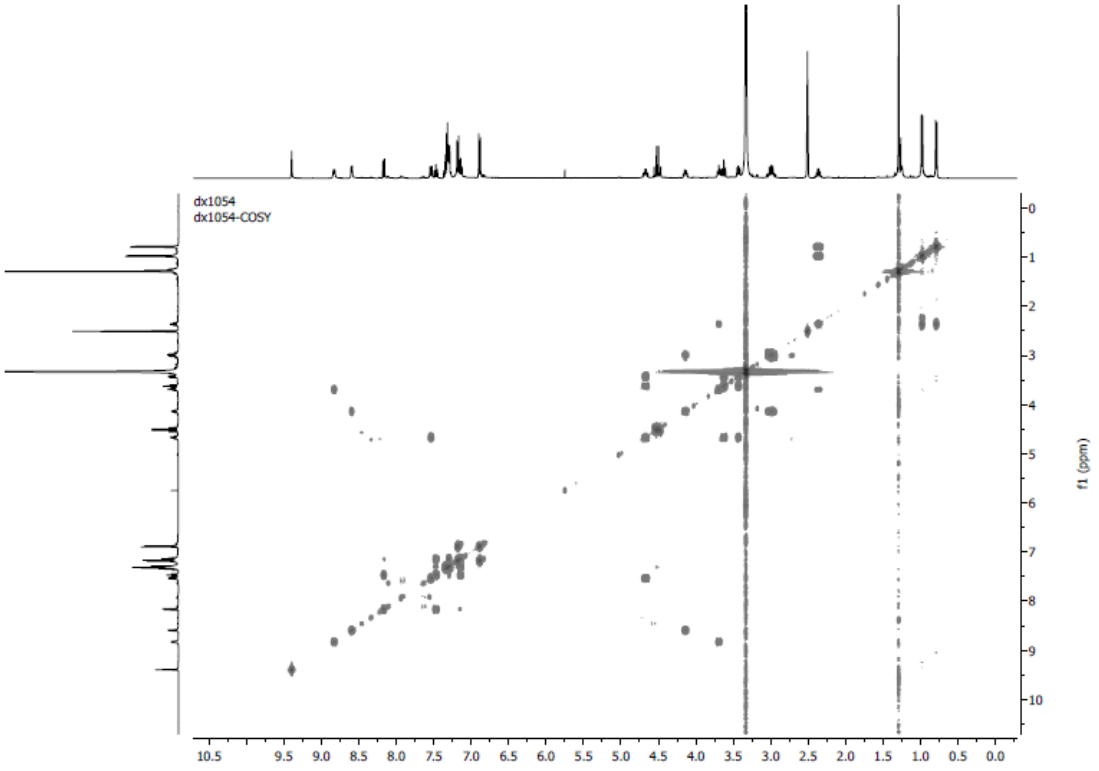
^{1}H NMR



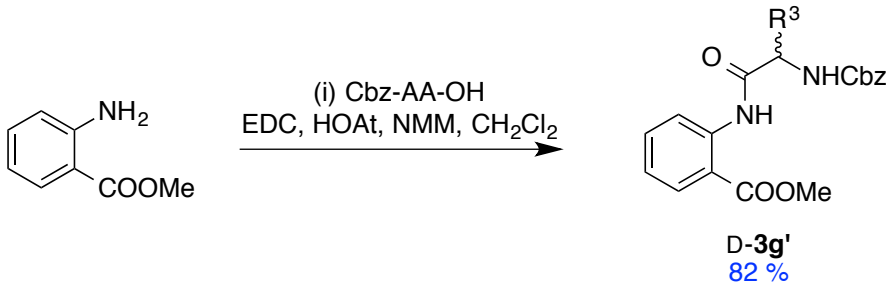
^{13}C NMR



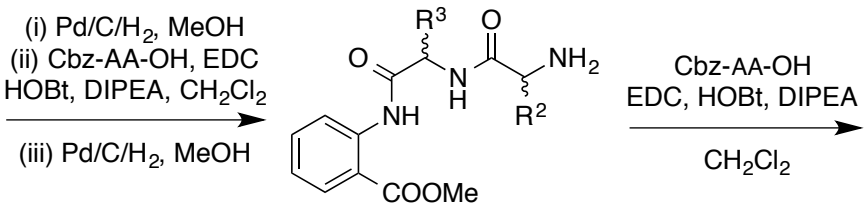
^{1}H-^{1}H NMR

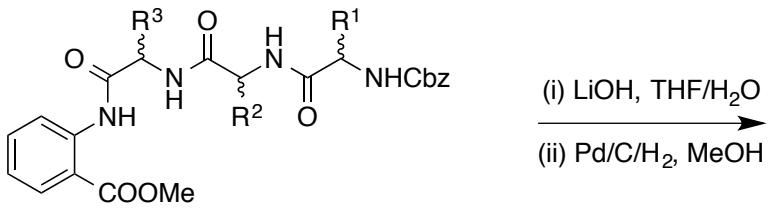

dx1054
dx1054-COSY
f1 (ppm)
10.5
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0

Cbz approach to products 1

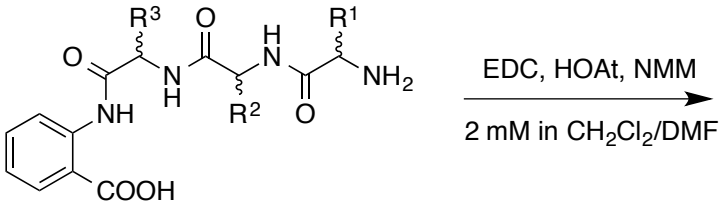


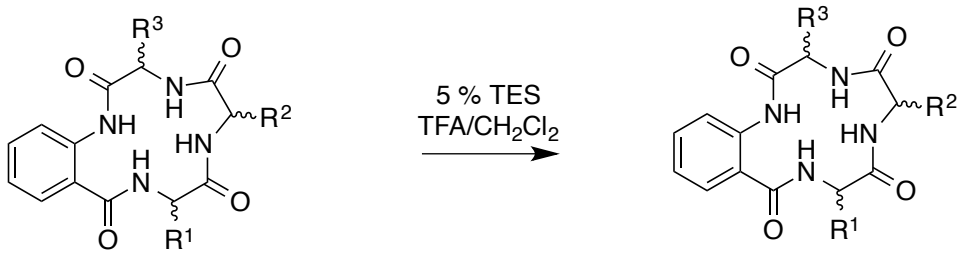
D-**3g'**
82 %





82 - 93 % over four steps





cyclo-LPhe-DAla-LPhe-Anth (LDL-**1faf**), 45 %
cyclo-DPhe-LAla-LGlu'-Anth (DLL-**1fae'**), 42 %

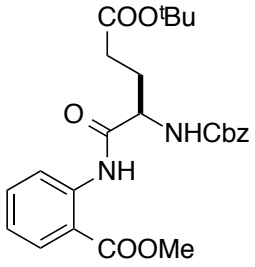
cyclo-DPhe-LAla-LGlu-Anth (DLL-**1fae**), 81 %

Scheme S2. Cbz approach to products **1**

General Procedures For Cbz Approach in Solution

Methyl anthranilate was coupled to the first Cbz protected amino acid with "Coupling Method I". The Cbz protection group was deprotected with "Cbz Deprotection Procedure" and the resulting material was coupled with the second Cbz protected amino acid with "Coupling Method II". The tripeptide intermediate was deprotected with "Cbz Deprotection Procedure" and coupled with the third Cbz amino acid with "Coupling Method II" to give the protected tetrapeptide intermediate. The methyl ester of this intermediate was hydrolyzed with the procedure "Hydrolysis of Methyl Ester" and the N- terminus Cbz group was removed with "Cbz Deprotection Procedure". The linear tetrapeptide was cyclized with "Procedure for Cyclization" to give the cyclic peptide product. If necessary, the cyclic peptide product was deprotected with "^{t}Bu Deprotection Procedure" to give the final deprotected product.

Methyl (R)-2-(2-(((benzyloxy)carbonyl)amino)-5-(tert-butoxy)-5-oxopentanamido)benzoate (D-**3g'**)

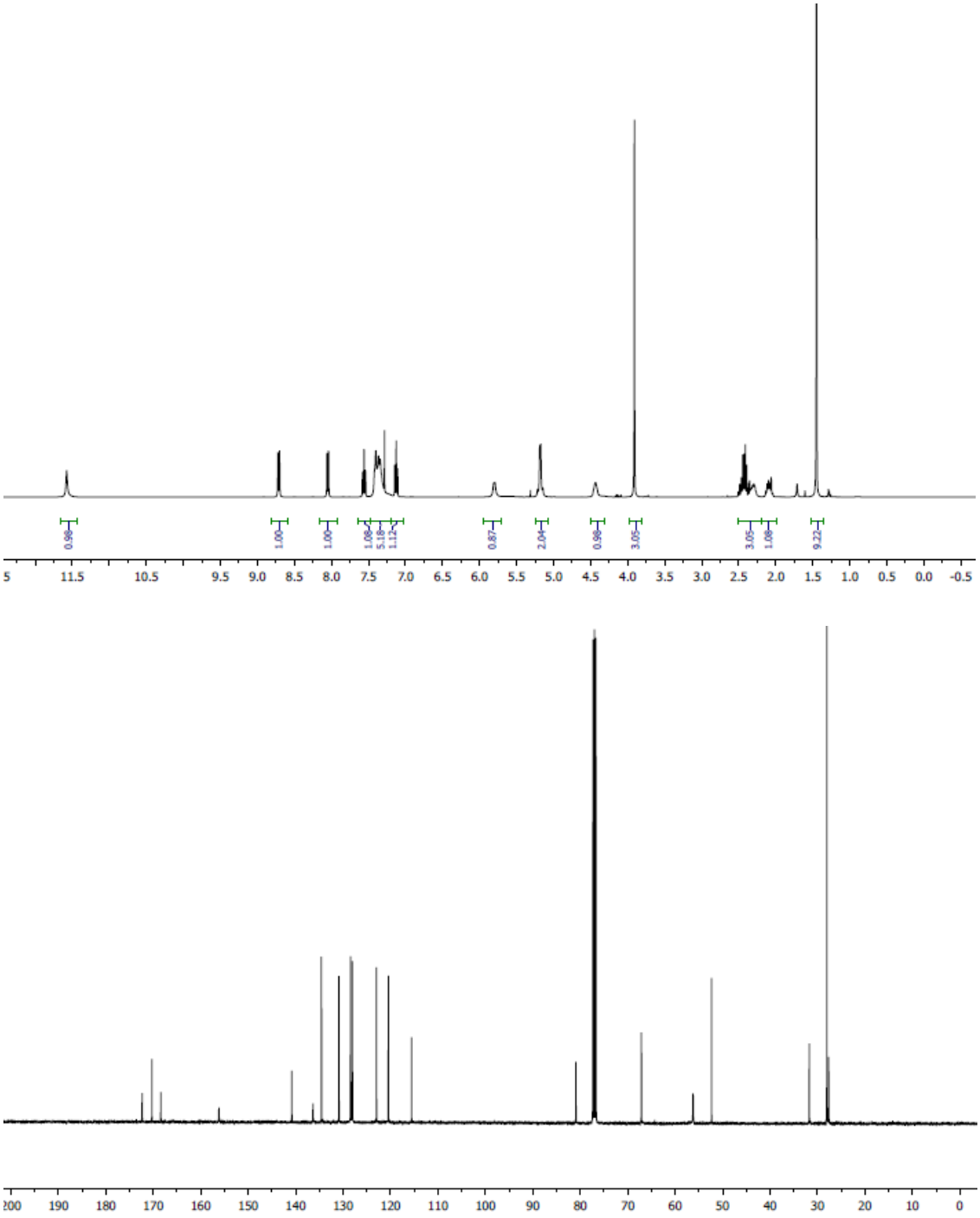


White solid, 82 %; mp = 78.8-79.8 °C;

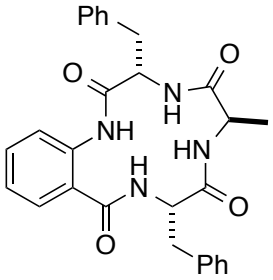
^{1}H-NMR (400 MHz, $CDCl_3$) δ 11.58 (s, 1H), 8.71 (d, J = 8.4 Hz, 1H), 8.05 (dd, J = 8.0, 1.5 Hz, 1H), 7.64-7.47 (m, 1H), 7.43-7.30 (m, 5H), 7.20-7.03 (m, 1H), 5.80 (d, J = 6.2 Hz, 1H), 5.28-5.10 (m, 2H), 4.49-4.38 (m, 1H), 3.91 (s, 3H), 2.50-2.19 (m, 3H), 2.19-1.99 (m, 1H), 1.45 (s, 9H);

^{13}C-NMR (101 MHz, $CDCl_3$) δ 172.3, 170.2, 168.4, 156.1, 140.8, 136.3, 134.5, 130.8, 128.4, 128.0, 128.0, 122.9, 120.4, 115.5, 80.9, 67.1, 56.2, 52.3, 31.7, 28.0, 27.6;

HRMS (ESI+) m/z calcd for $C_{25}H_{31}N_2O_7$ $(M+H)^+$ 471.2131; found 471.2145.



(3S,6R,9S)-3,9-Dibenzyl-6-methyl-3,4,6,7,9,10-hexahydro-1H-benzo[k][1,4,7,10]tetraazacyclotridecine-2,5,8,11-tetraone (LDL-**1faf**)



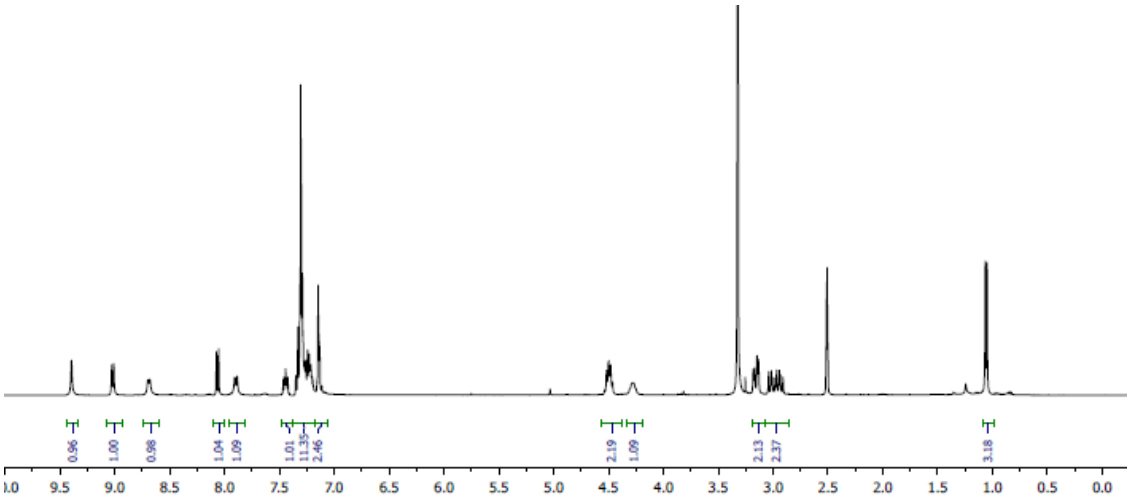
White solid, 45 % over four steps, starting from Boc-deprotected L-**2f**; mp > 270 °C;

^{1}H-NMR (400 MHz, DMSO-d_6) δ 9.39 (s, 1H), 9.02 (d, J = 7.5 Hz, 1H), 8.69 (d, J = 7.1 Hz, 1H), 8.06 (d, J = 8.2 Hz, 1H), 7.90 (d, J = 8.8 Hz, 1H), 7.49-7.37 (m, 1H), 7.37-7.17 (m, 10H), 7.17-7.06 (m, 2H), 4.53-4.48 (m, 2H), 4.29-4.25 (m, 1H), 3.19-3.12 (m, 2H), 3.08-2.90 (m, 2H), 1.06 (d, J = 6.8 Hz, 3H);

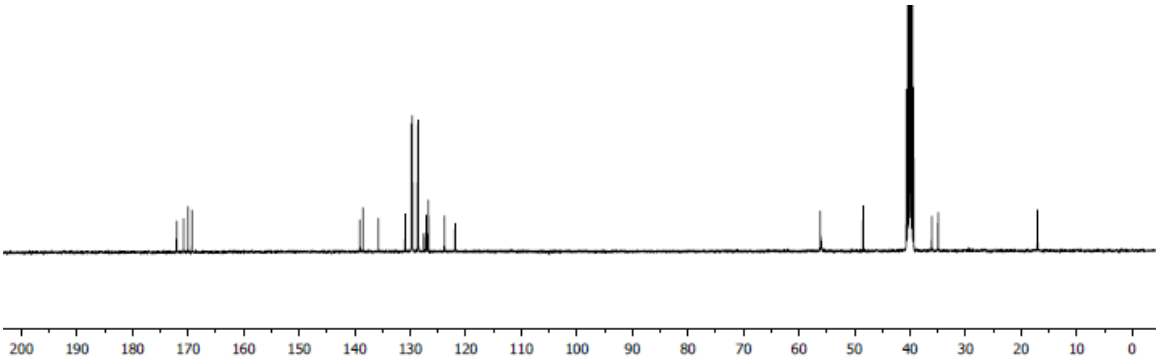
^{13}C-NMR (101 MHz, DMSO-d_6) δ 172.0, 170.8, 170.0, 169.2, 139.0, 138.5, 135.7, 130.9, 129.7, 129.6, 128.6, 128.5, 127.5, 127.1, 126.7, 126.7, 123.8, 121.9, 56.1, 56.0, 48.4, 36.0, 34.9, 17.0;

HRMS (ESI+) m/z calcd for $C_{28}H_{28}N_4O_4Na$ $(M+Na)^+$ 507.2008; found 507.2027.

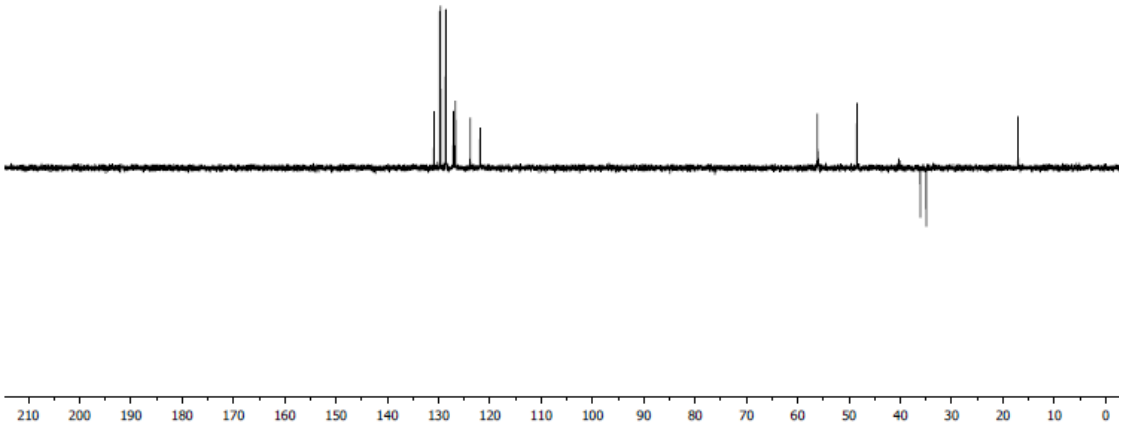
^{1}H NMR



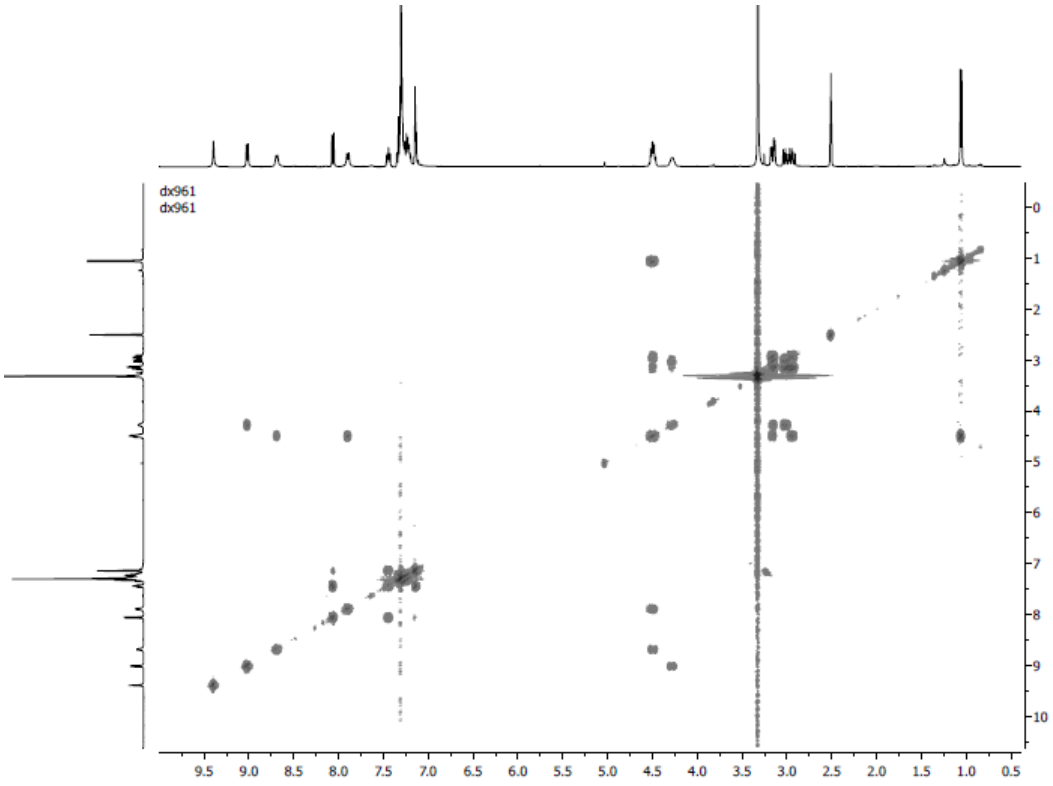
^{13}C NMR



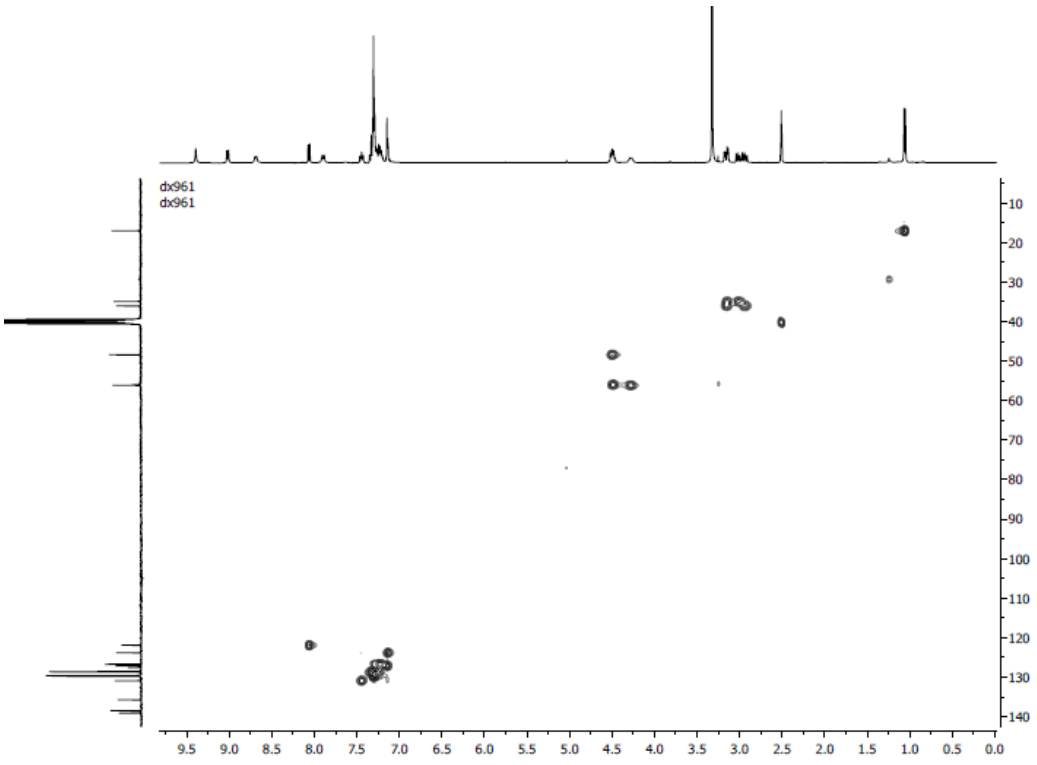
DEPT135



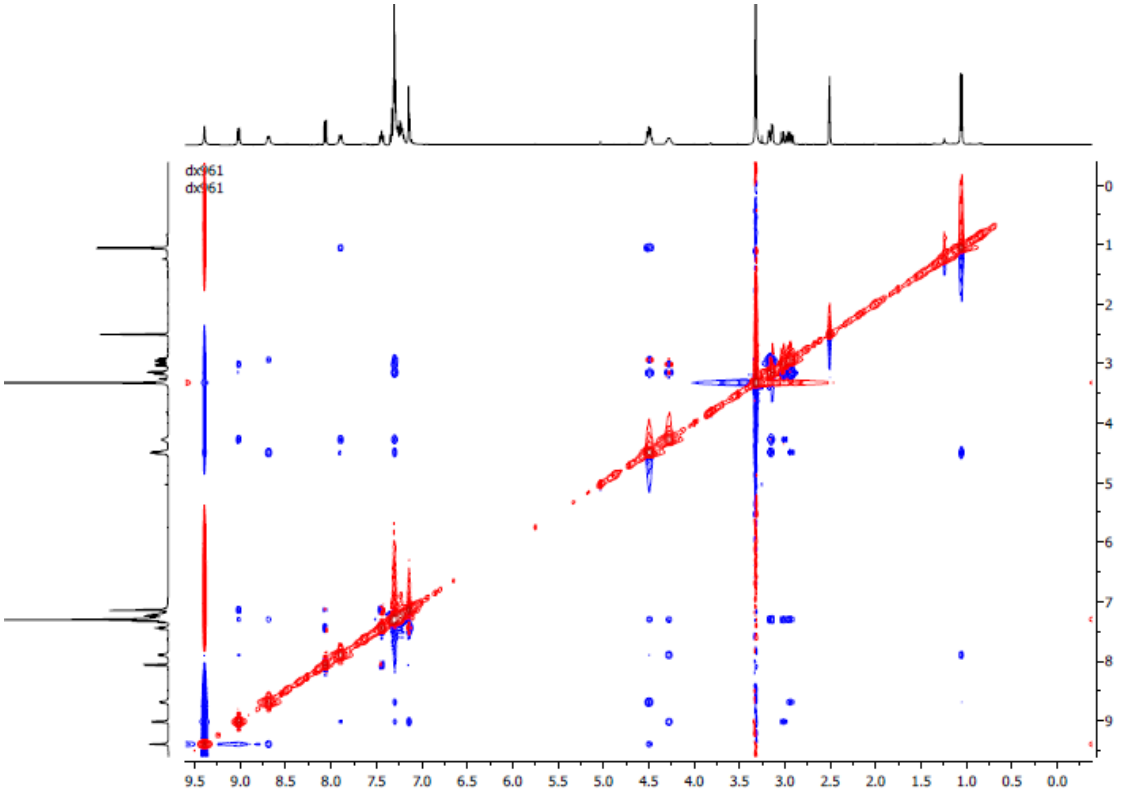
^{1}H-^{1}H COSY



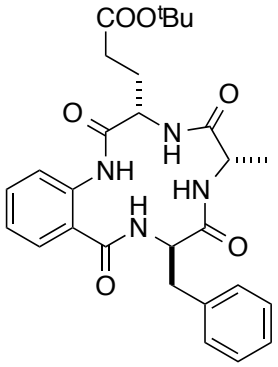
^{1}H-^{13}C HSQC


dx961
dx961
10
20
30
40
50
60
70
80
90
100
110
120
130
140
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0

NOESY


dx961
dx961
0
1
2
3
4
5
6
7
8
9
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0

tert-Butyl 3-((3S,6S,9R)-9-benzyl-6-methyl-2,5,8,11-tetraoxo-2,3,4,5,6,7,8,9,10,11-decahydro-1H-benzo[k][1,4,7,10]tetraazacyclotridecin-3-yl)propanoate (DLL-**fae'**)



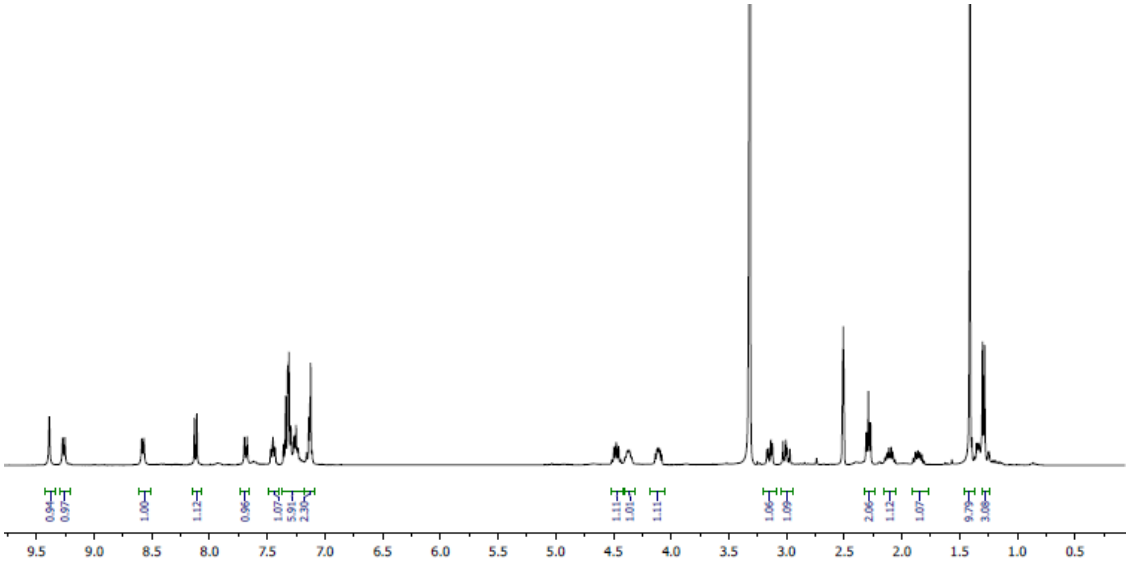
White solid, 42 % over three steps; mp > 270 °C, color change observed at above 215 °C;

^{1}H-NMR (400 MHz, DMSO-d_6) δ 9.39 (s, 1H), 9.26 (d, J = 7.9 Hz, 1H), 8.58 (d, J = 7.4 Hz, 1H), 8.12 (d, J = 8.2 Hz, 1H), 7.69 (d, J = 9.2 Hz, 1H), 7.49-7.40 (m, 1H), 7.37-7.18 (m, 5H), 7.18-7.08 (m, 2H), 4.52-4.42 (m, 1H), 4.41-4.32 (m, 1H), 4.16-4.08 (m, 1H), 3.15 (dd, J = 13.8, 5.4 Hz, 1H), 3.00 (dd, J = 13.8, 9.7 Hz, 1H), 2.29 (t, J = 7.5 Hz, 2H), 2.15-2.05 (m, 1H), 1.91-1.77 (m, 1H), 1.41 (s, 9H), 1.29 (d, J = 7.2 Hz, 3H);

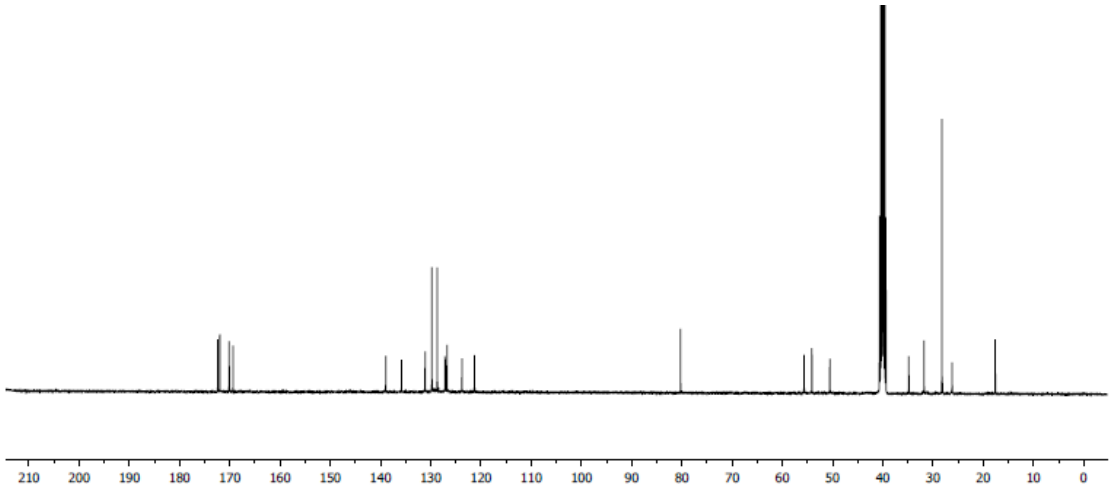
^{13}C-NMR (101 MHz, DMSO-d_6) δ 172.4, 172.0, 170.1, 170.1, 169.3, 138.9, 135.8, 131.1, 129.7, 128.7, 127.1, 127.0, 126.7, 123.7, 121.3, 80.2, 55.6, 54.1, 50.5, 34.8, 31.7, 28.2, 26.2, 17.6;

HRMS (ESI-) m/z calcd for $C_{28}H_{33}N_4O_6$ $(M-H)^-$ 521.2400; found 521.2411.

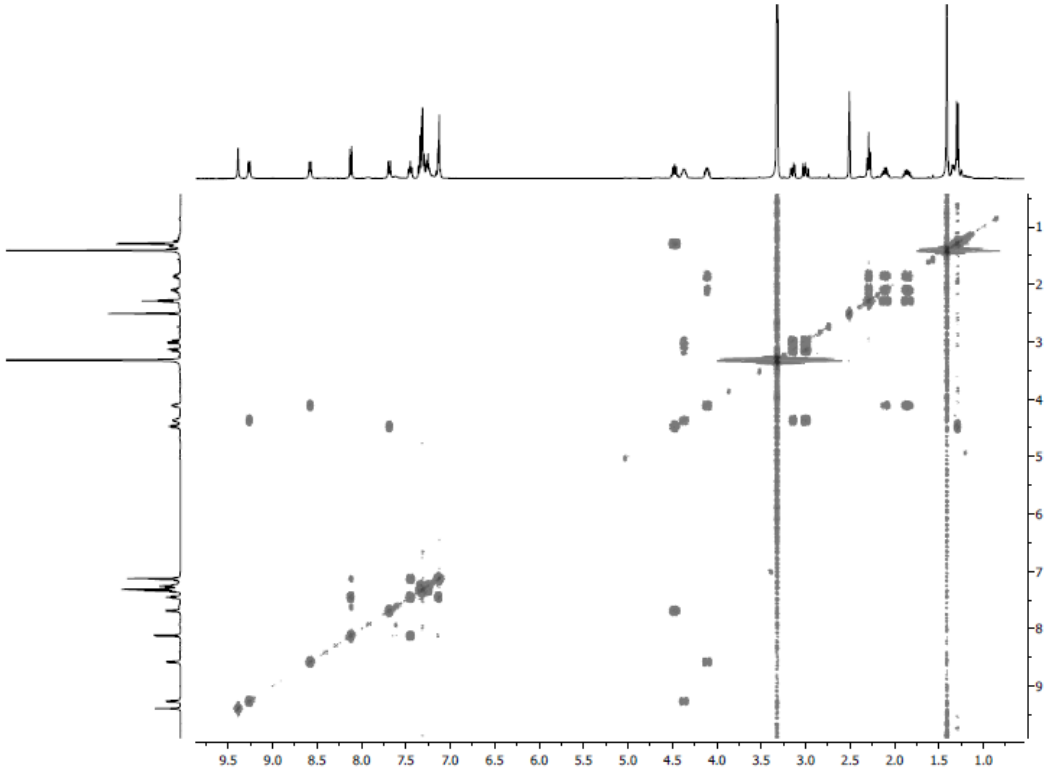
^{1}H NMR


0.94
0.97
1.00
1.12
0.96
1.07
5.91
2.30
1.11
1.01
1.11
1.06
1.09
2.06
1.12
1.07
9.79
3.08
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5

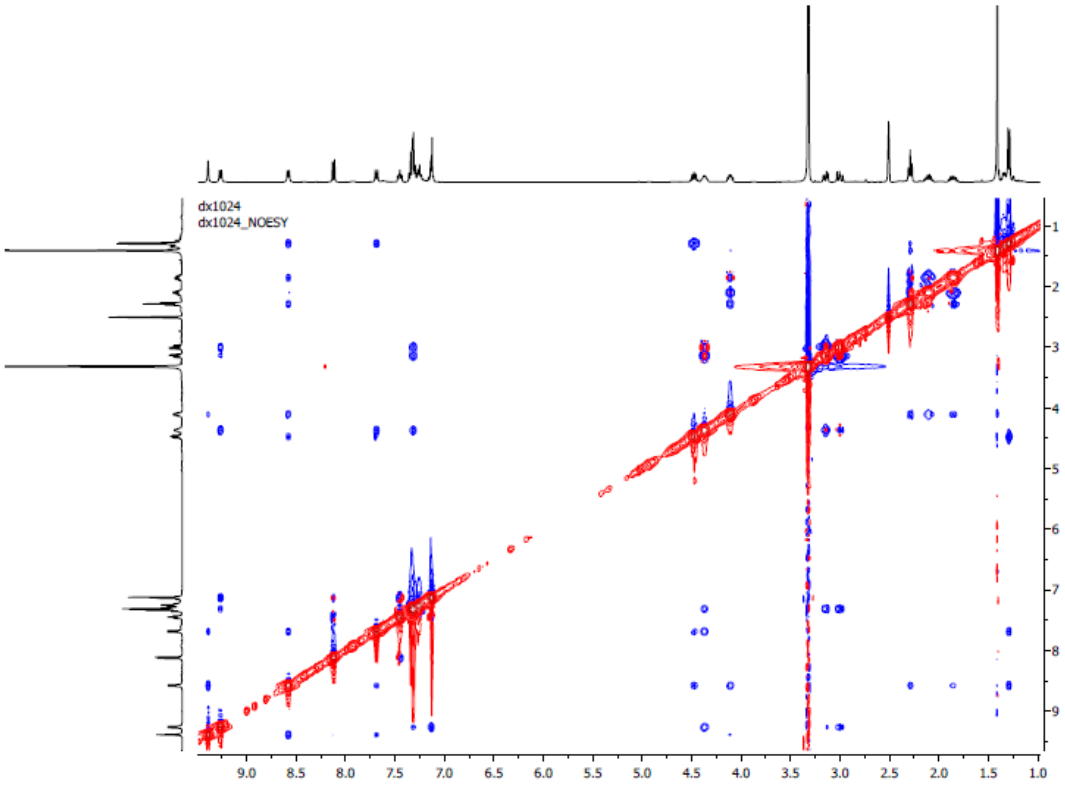
^{13}C NMR


210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0

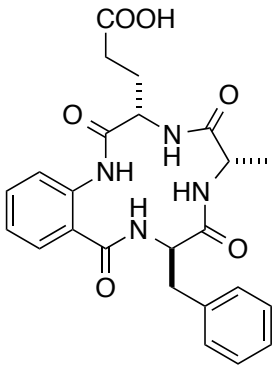
^{1}H-^{1}H COSY



NOESY



3-((3S,6S,9R)-9-Benzyl-6-methyl-2,5,8,11-tetraoxo-2,3,4,5,6,7,8,9,10,11-decahydro-1H-benzo[k][1,4,7,10]tetraazacyclotridecin-3-yl)propanoic acid (DLL-**fae**)



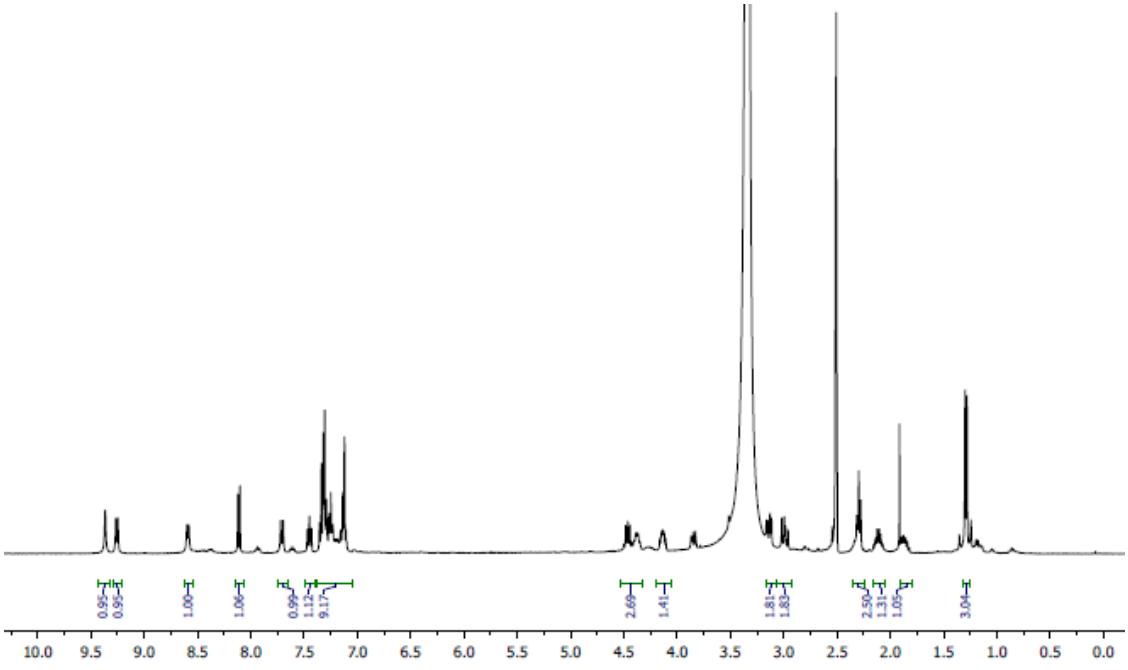
White solid, 81 %; mp > 270 °C;

^{1}H-NMR (400 MHz, DMSO-d_6) δ 9.37 (s, 1H), 9.26 (d, J = 8.0 Hz, 1H), 8.59 (d, J = 7.3 Hz, 1H), 8.11 (d, J = 8.1 Hz, 1H), 7.71 (d, J = 9.1 Hz, 1H), 7.50-7.39 (m, 1H), 7.37-7.05 (m, 7H), 4.54-4.32 (m, 2H), 4.19-4.06 (m, 1H), 3.14 (dd, J = 13.8, 5.4 Hz, 1H), 2.99 (dd, J = 13.8, 9.6 Hz, 1H), 2.30 (t, J = 7.6 Hz, 2H), 2.15-2.08 (m, 1H), 1.90-1.79 (m, 1H), 1.29 (d, J = 7.3 Hz, 3H);

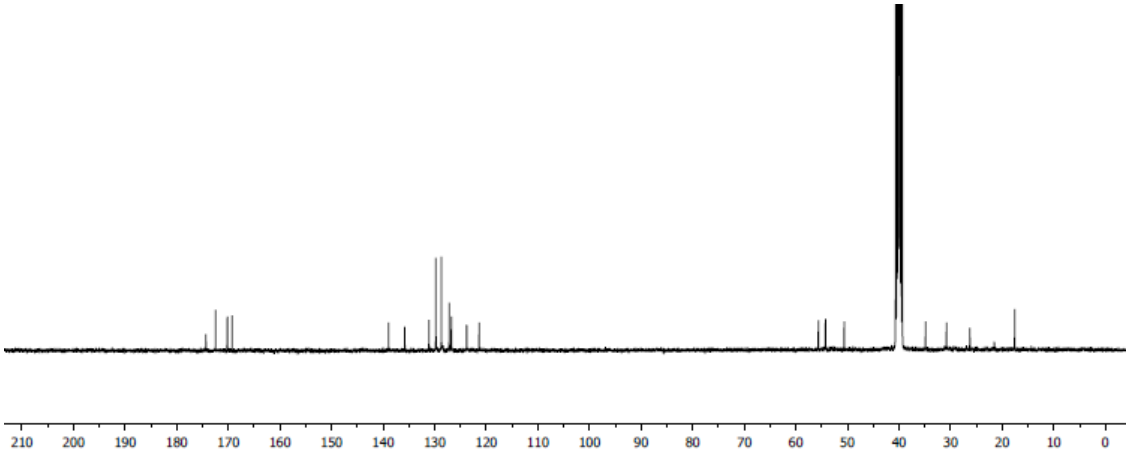
^{13}C-NMR (101 MHz, DMSO-d_6) δ 174.3, 172.4, 170.2, 170.1, 169.2, 138.9, 135.8, 131.1, 129.7, 128.7, 127.1, 127.0, 126.7, 123.8, 121.3, 55.6, 54.2, 50.6, 34.8, 30.8, 26.2, 17.5;

HRMS (ESI-) m/z calcd for $C_{24}H_{25}N_4O_6$ $(M-H)^-$ 465.1774; found 465.1781.

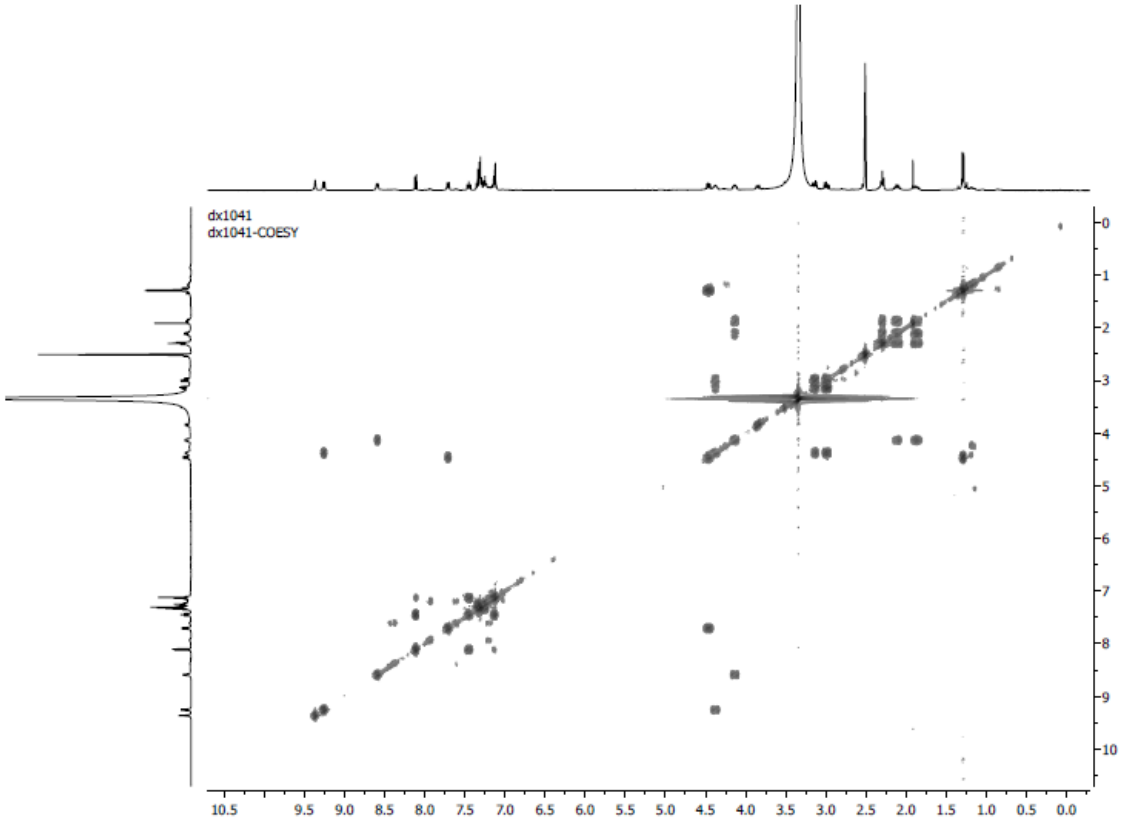
^{1}H NMR



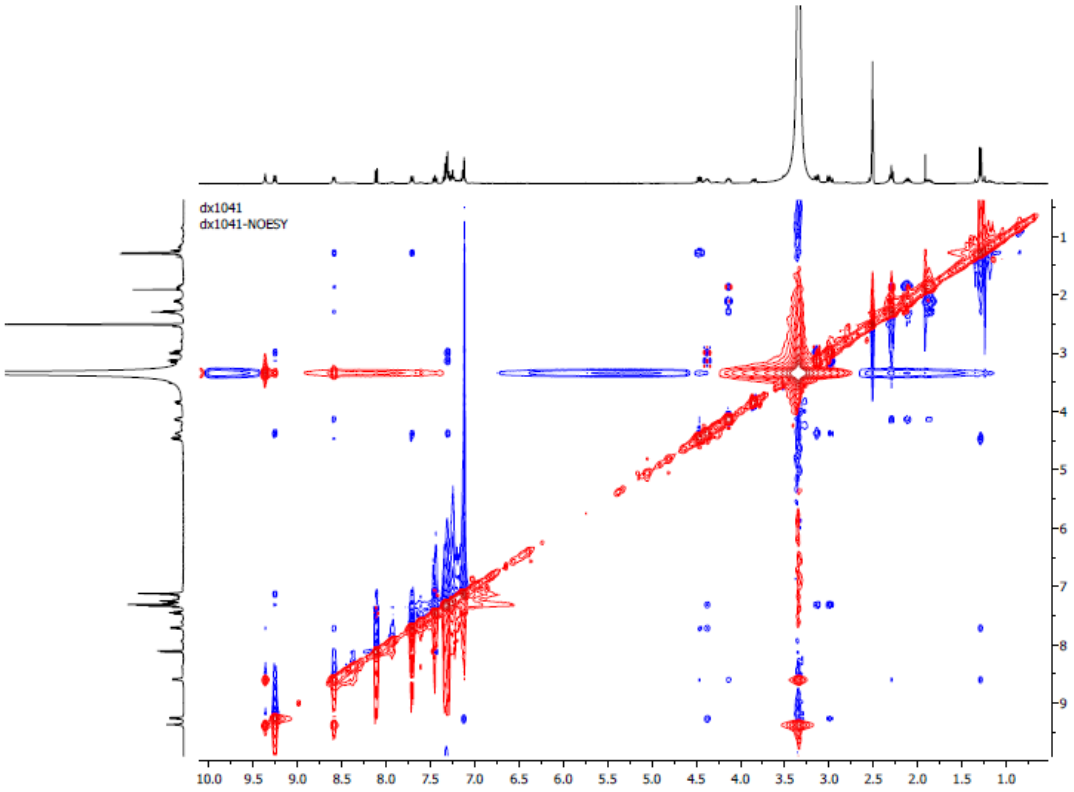
^{13}C NMR


210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0

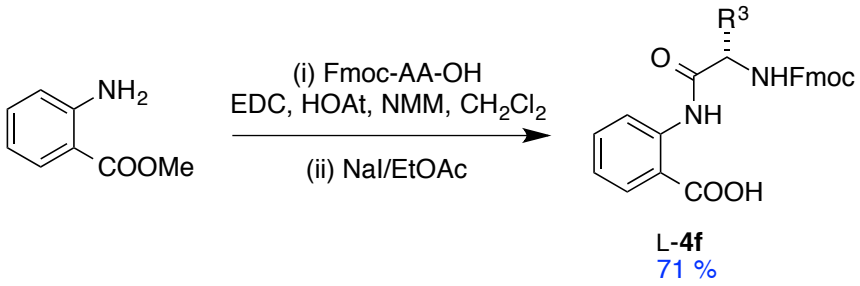
^{1}H-^{1}H COSY


dx1041
dx1041-COESY
10.5
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
0
1
2
3
4
5
6
7
8
9
10

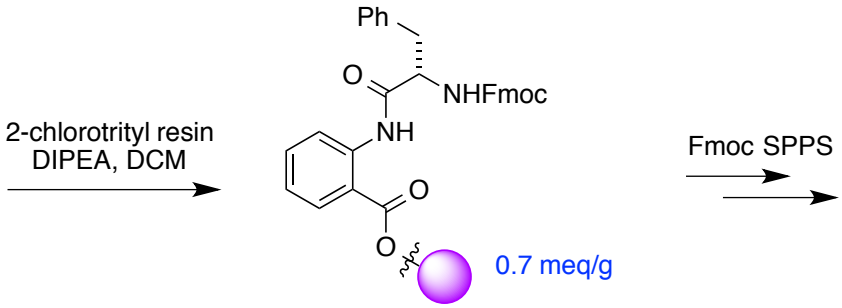
NOESY


dx1041
dx1041-NOESY
10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0
1 2 3 4 5 6 7 8 9

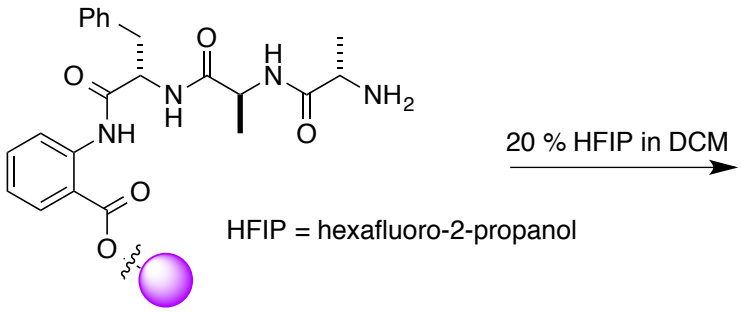
Fmoc solid phase approach to products 1



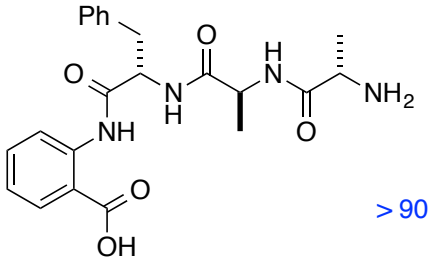
L-**4f**
71 %



0.7 meq/g



HFIP = hexafluoro-2-propanol



89 % yield
> 90 % pure by ^{1}H NMR

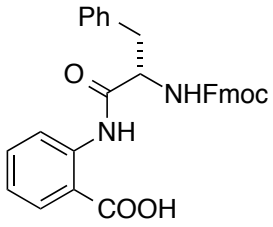
Scheme S3. Fmoc solid phase approach to products **1.**

General Procedures for Fmoc Approach on Solid Phase

Synthesis Of Dipeptide Intermediate

Fmoc-Phe-OH (775 mg, 2.0 mmol, 1.0 equiv.) was added to a solution of HOAt (408 mg, 3.0 mmol, 1.5 equiv.) and NMM (220 μL, 2.0 mmol, 1.0 equiv.) in DCM (3.0 mL) at 0 ^{o}C under N_2. The resulting mixture was stirred at 0 ^{o}C for 10 min and then methyl anthranilate (906 mg, 6.0 mmol, 3.0 equiv.) was added, followed by the addition of EDC•HCl (575 mg, 3.0 mmol, 1.5 equiv.) in one portion. The reaction mixture was allowed to warm to room temperature over 1 h and was stirred at room temperature for 2 h. 30 mL DCM was added to dilute the solution and the organic phase was washed with 0.2 M HCl aqueous solution (30 mL × 5). The organic phase was further washed with brine (10 mL), saturated $NaHCO_3$ solution (20 mL × 2) and brine (10 mL), dried over $MgSO_4$ and filtered. The solvent was removed under vacuum to give the crude material. The crude material was dissolved in EtOAc (10 mL) and LiI (1.07 g, 8.0 mmol, 4.0 equiv.) was added. The mixture was stirred at 80 ^{o}C for 18 h and then it was cooled down to room temperature. 0.2 M HCl (20 mL) was added and the mixture was extracted with EtOAc (30 mL × 3). The organic phase was washed with brine (10 mL), dried over $MgSO_4$, filtered and the solvent was removed under vacuum to give the crude product. The crude product was purified by flash chromatography (30% EtOAc in DCM to 100 % EtOAc) to give the desired product L-**4f** as a white solid.

(S)-2-(2-((((9H-Fluoren-9-yl)methoxy)carbonyl)amino)-3-phenylpropanamido)benzoic acid (L-**4f**)



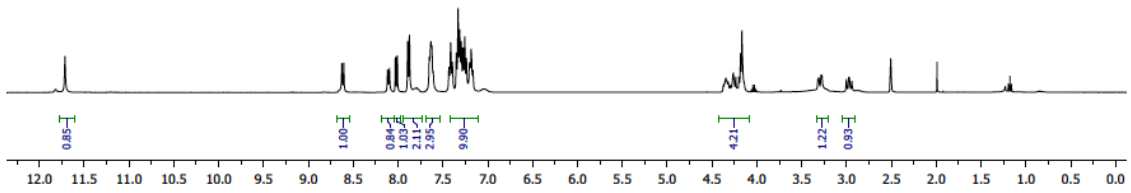
White solid, 71 %; mp = 175.5-176.4 °C;

^{1}H-NMR (400 MHz, DMSO-d_6) δ 11.71 (s, 1H), 8.62 (d, J = 8.4 Hz, 1H), 8.11 (d, J = 8.0 Hz, 1H), 8.02 (dd, J = 7.9, 1.4 Hz, 1H), 7.88-7.83 (m, 2H), 7.68-7.57 (m, 3H), 7.42-7.11 (m, 10H), 4.43-4.08 (m, 4H), 3.30 (dd, J = 13.8, 4.2 Hz, 1H), 2.96 (dd, J = 13.6 Hz, 10.7 Hz, 1H);

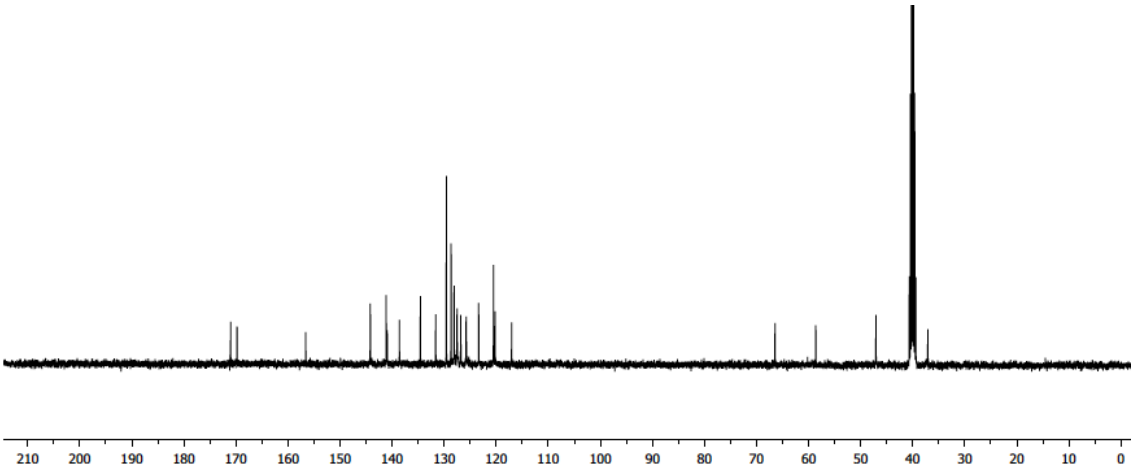
^{13}C-NMR (101 MHz, DMSO-d_6) δ 171.0, 169.8, 156.6, 144.2, 141.1, 140.9, 138.6, 134.6, 131.6, 129.6, 128.7, 128.1, 127.5, 126.8, 125.8, 123.4, 120.5, 120.2, 117.1, 66.5, 58.6, 47.1, 37.1;

HRMS (ESI-) m/z calcd for $C_{31}H_{25}N_2O_5$ $(M-H)^-$ 505.1763; found 505.1781.

^{1}H NMR



^{13}C NMR



Loading of Dipeptide onto 2-Cl-Trityl Resin

Cl-Trt resin (200 mg, 1.4 meq/g) was shaken with anhydrous DCM (4 mL) in a fritted syringe for 30 min. Then the DCM was removed and a mixture of L-**4f** (71 mg, 0.14 mmol) and DIPEA (98 μL, 0.56 mmol) in DCM (2 mL) was added into the syringe and the mixture was shaken at room temperature for 2h. The remaining reactive site was blocked with MeOH/DIPEA (9:1 v/v) for 30 min and the beads were washed with DCM 3 times, MeOH and then DMF 3 times.

Coupling With Amino Acids And Fmoc Deprotection

Fmoc protection groups were deprotected by treating the bead with 20 % piperidine in DMF for 1 min, followed by the second treatment with 20 % piperidine in

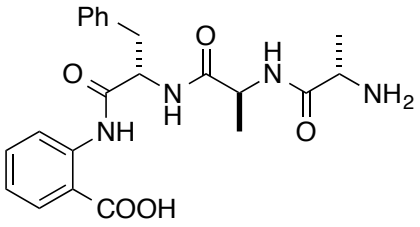
DMF for 15 minutes. The beads were washed with DMF 6 times after the second treatment.

Coupling reactions with amino acids were carried out with 3 equiv. of Fmoc amino acid, 3 equiv. of HBTU, 3 equiv. of HOBt, 6 equiv. of DIPEA in DMF for 1 h at room temperature. The beads were washed with DMF 6 times after the coupling reaction and a few beads were subjected to Kaiser test to confirm the completion of the coupling reaction.

Cleavage From Solid Support

After the last Fmoc deprotection step, the resin was washed with DMF 6 times, MeOH 3 times and DCM 3 times. The linear peptide was cleaved off the bead by treating the beads with HFIP/DCM (1:4 v/v) for 30 min at room temperature. After filteration, the solvents were removed under vacuum and the crude material was dried under high vacuum to give the linear peptide. The crude material was analyzed by HPLC and ^{1}H NMR for its purity.

2-((S)-2-((S)-2-((S)-2-Aminopropanamido)propanamido)-3-phenylpropanamido)benzoic acid (LLL-**aaf** linear peptide, crude material)



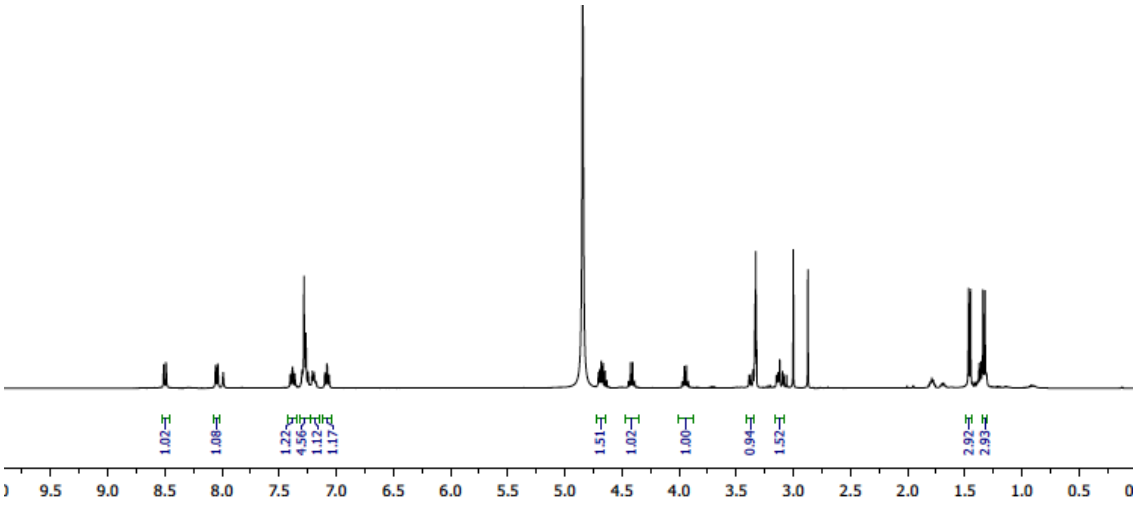
White solid, 89 %;

^{1}H-NMR (400 MHz, MeOH-d_4) δ 8.50 (dd, J = 8.3, 0.9 Hz, 1H), 8.05 (dd, J = 7.8, 1.6 Hz, 1H), 7.43-7.35 (m, 1H), 7.32-7.23 (m, 4H), 7.23-7.17 (m, 1H), 7.08 (td, J = 7.8, 1.1 Hz, 1H), 4.72-4.65 (m, 1H), 4.42 (q, J = 7.1 Hz, 1H), 3.95 (q, J = 7.0 Hz, 1H), 3.41-3.34 (m, 1H), 3.16-3.08 (m, 1H), 1.46 (d, J = 7.0 Hz, 3H), 1.33 (d, J = 7.1 Hz, 3H).

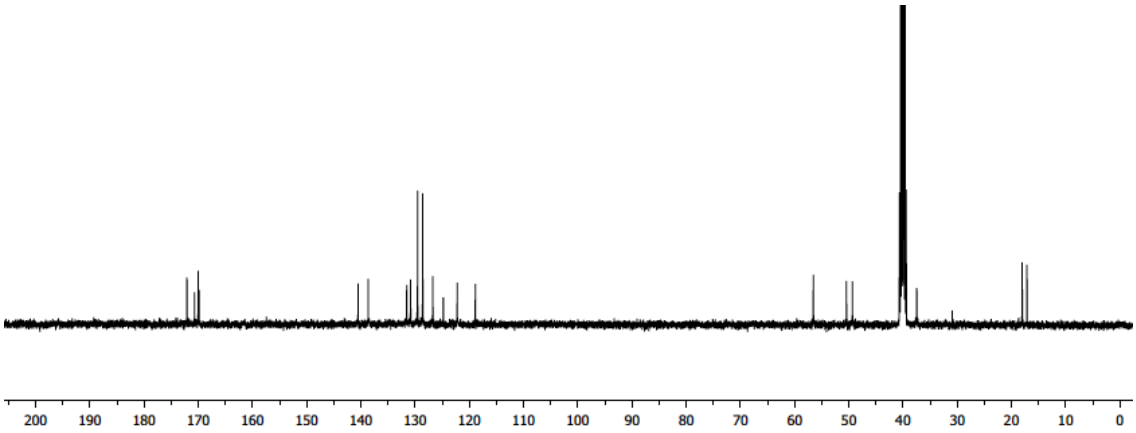
^{13}C-NMR (101 MHz, DMSO-d_6) δ 172.0, 170.6, 170.0, 169.8, 140.5, 138.6, 131.5, 130.8, 129.5, 128.5, 126.7, 124.7, 122.2, 118.8, 56.5, 50.4, 49.3, 37.4, 18.0, 17.1;

HRMS (ESI+) m/z calcd for $C_{22}H_{27}N_4O_5$ $(M+H)^+$ 427.1981; found 427.1999.

^{1}H NMR



^{13}C NMR

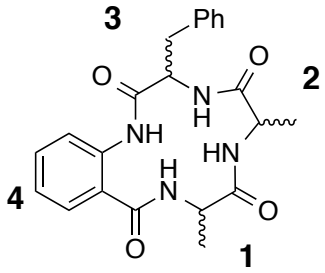


B. NMR Structure Determination and Analysis

NMR measurements for cyclic peptides **1** were carried out in DMSO-d_6 with a sample concentration of ~20 mM. NOESY spectra were taken using a mixing time of 400 ms for compounds LLL-**1aaf**, LDL-**1aaf**, DLL-**1aaf**, DDL-**1aaf**, DDL-**1e'af**, DDL-**1eaf**, LLL-**1vsy**, DLL-**1fae'**, DLL-**1fae**, LDL-**1faf**. ROESY spectra were taken using a mixture time of 200 ms for compounds LLL-**1vs'y'** and DDD-**1y's'v**. There was no evidence of cis-amide bonds due to the absence of Cα-Cα or Cα-Cβ couplings across residues.

The observed NOE measurements were summarized in the following tables for **1aaf**. s: strong, 1.8 Å ≤ H-H distance ≤ 2.7 Å; m: medium, 1.8 Å ≤ H-H distance ≤ 3.5 Å; w: weak, 1.8 Å ≤ H-H distance ≤ 5.0 Å.[20,21]

Table S1. Observed NOE measurements.



LLL-**1aaf**

cross peak		intensity	cross peak		intensity
				$H\beta^1$	m
	$H\beta^2$	m		$H\beta^2$	m
	$H\beta^3$	w		$H\alpha^1$	w
NH^3	$H\alpha^2$	m	NH^2	$H\alpha^2$	w
	$H\alpha^2$	m		NH^1	m
	NH^4	m		NH^3	w
				NH^4	w

NH^1	$H\beta^1$	m	NH^4	$H\alpha^3$	w
	$H\alpha^1$	w			

LDL-**1aaf**

cross peak		intensity	cross peak		intensity
	$H\beta^2$	w		$H\beta^2$	m
	$H\beta^3$	m		$H\alpha^1$	m
NH^3	$H\alpha^2/H\alpha^3$	s	NH^2	$H\alpha^2$	w
	NH^2	w		NH^1	w
	NH^4	m		NH^3	w
				NH^4	w
NH^1	$H\beta^1$	m	NH^4	$H\beta^3$	w
	$H\alpha^1$	m		$H\alpha^3$	m

DLL-**1aaf**

cross peak		intensity	cross peak		intensity
	$H\beta^2$	m		$H\beta^2$	m
	$H\beta^3$	m		$H\alpha^1 + H\alpha^2$	s
NH^3	$H\alpha^2 + H\alpha^3$	m	NH^2	NH^1	w
	NH^2	w		NH^4	w
	NH^4	m			
NH^1	$H\beta^1$	m	NH^4	$H\alpha^3$	w
	$H\alpha^1$	m			

DDL-**1aaf**

cross peak		intensity	cross peak		intensity
	$H\beta^3$	m		$H\beta^2$	m
	$H\alpha^3$	w		$H\beta^1$	m
NH^3	$H\alpha^2$	s	NH^2	$H\alpha^1$	w
	NH^4	m		$H\alpha^2$	w
				NH^1	w

NH^1	$H\beta^1$	s	NH^4	$H\beta^3$	w
	$H\alpha^1$	w		$H\alpha^3$	w

NMR structure calculation in DMSO was carried out using the Conformational Searches in MacroModel with distance checks (MacroModel, version 10.0, Schrödinger, LLC, New York, NY). Monte Carlo Multiple Minimum method were used to sample the cyclic peptide conformations. 10000 structures were sampled and minimized with OPLS_2005 force field in dielectric constant 46.7 with a convergence criteria 0.05 kJ/mol over 2000 iterations. Distance constraints from the previous tables were applied during the conformational sampling to eliminate the conformations with distance violations. Duplicate structures base on heavy-atom superposition (RMSD < 0.02 Å) were discarded. The unique conformations within 5 kJ/mol of the global minimum were collected and clustered based on heavy atoms of the macrocyclic scaffold without the Phe side-chain. The clustered conformations were shown in Figure S4.

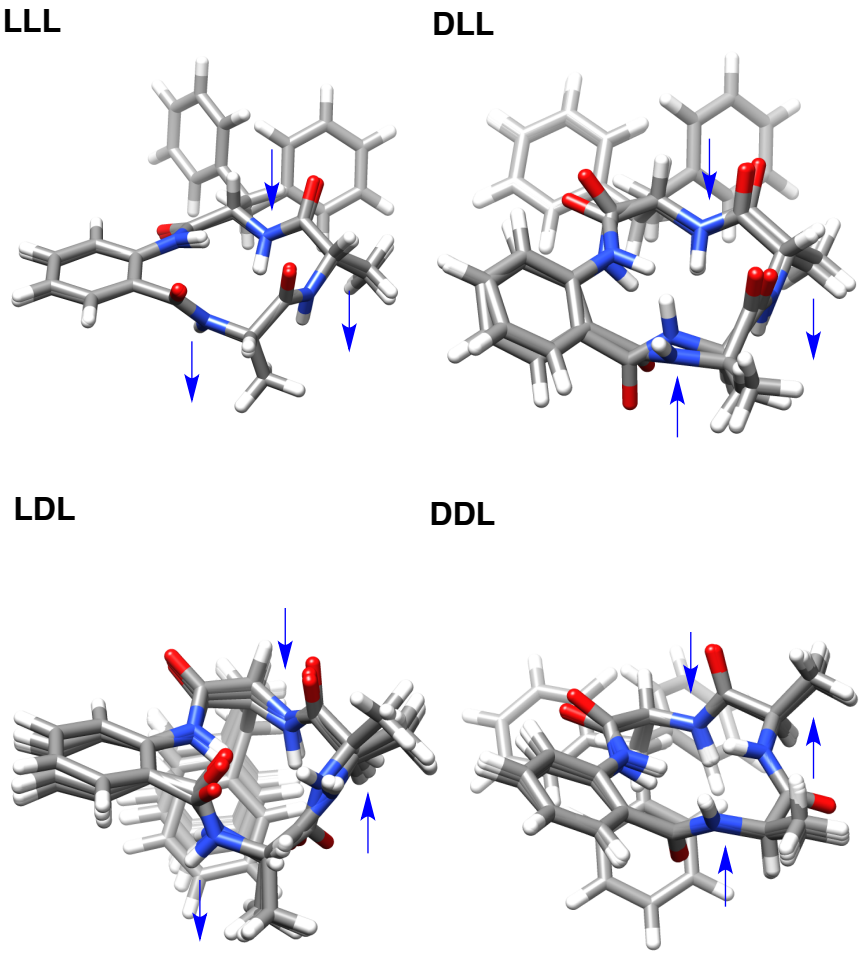


Figure S4. Conformational clusters of **1aaf** with NMR constraints.

C. H/D Exchange of the Amide NH

H/D exchange experiments were explored in $CDCl_3/CD_3OD$. A sample of cyclic peptide **1** was prepared in 450 μL $CDCl_3$ and 50 μL of CD_3OD was added (4 mM final cyclic peptide concentration). The mixture was mixed for 1 min and ^1H NMR of the sample was recorded 3, 5, 7, 10, 15, 20, 25, 30 min after the addition of CD_3OD. The fact that all the amide protons showed significant reduction (> 75 %) in their intensity 30 min after CD_3OD addition indicated that there was no strong intramolecular hydrogen-bond interaction within any diastereomer of the cyclic peptide under the conditions tested.[22,23] The half-lifes of the H/D exchange reactions were summarized in Table S2.

Table S2. H/D exchange half-life $t_{1/2}$ (min).

	NH[1]	**NH[2]**	**NH[3]**	**NH[4]**
LLL-**1aaf**	< 5	15	< 5	< 5
DLL-**1aaf**	< 5	12	< 5	< 5
LDL-**1aaf**	< 5	two peaks overlapped*		< 5
DDL-**1aaf**	< 5	< 5	< 5	< 5

***** For the overlapped peak, the peak intensity decreased to 19 % of the original intensity after 30 min.

D. Temperature Dependence of the Amide NH Chemical Shifts

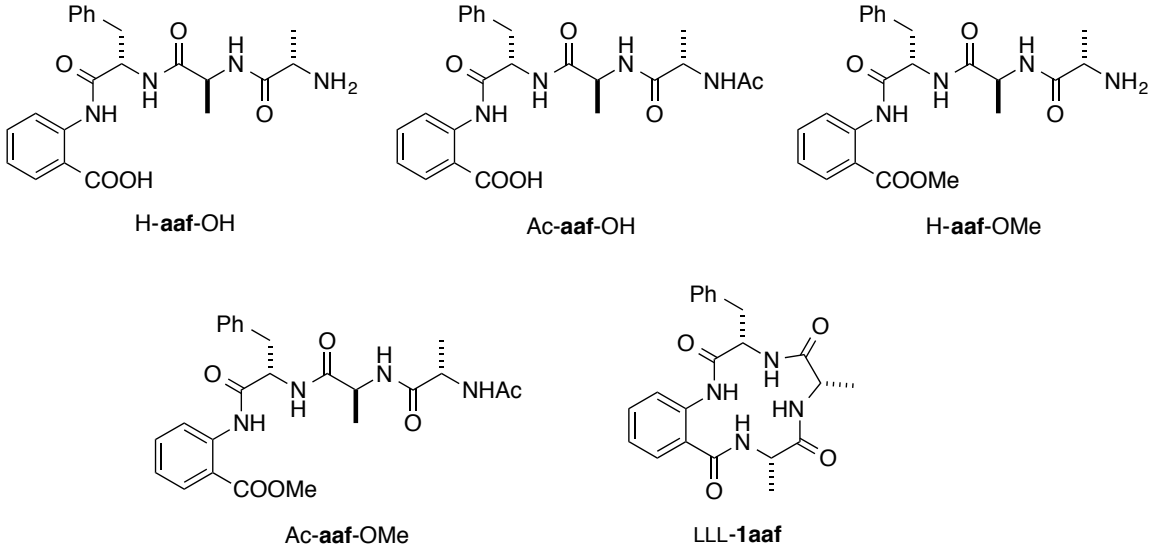
A ~10 mM solution of **1** in DMSO-d_6 was prepared. ^{1}H NMR measurements were made in the range 303 – 353 K. The first measurement was made at 303 K and the rest of ^{1}H NMRs were acquired at 10 K intervals. All the spectra obtained were referenced to the solvent peak and the ppm change of the amide NH peaks was monitored. The change in chemical shift was plotted versus the change in temperature and the data was fitted to a linear equation to give the temperature-dependent coefficient ($\Delta\delta/\Delta K$) of the NH proton of interest. The temperature coefficient data in Table S3 indicated that most of the amide protons were shielded from the solvent in the cyclic peptide system.[24,25]

Table S3. Temperature coefficient data of amide NHs.

	NH^1	**NH^2**	**NH^3**	**NH^4**
LLL-**1aaf**	-1.7	-	-4.5	-2.7
DLL-**1aaf**	-4.2	-1.6	-4.9	-2.7
LDL-**1aaf**	-3.7	-3.8	-6.2	-3.7
DDL-**1aaf**	-2.3	-	-6.8	-3.4

E. QikProp Calculation

QikProp 3.5 from Schrödinger (2012)[26] was used to evaluate pharmaceutically relevant properties for compounds listed below.



F. Stability Analysis of Cyclic Peptide

General Procedure for the pH Stability Assay

A stock solution of cyclic peptide LLL-**1aaf** (66 mM) and triphenylphosphine oxide (TPPO, internal standard, 100 mM) in DMSO was prepared and stored at 25 °C. 9 μL of the DMSO stock solution was dissolved in aqueous solutions with different pH (pH 7.4, PBS buffer; pH 12, 10 mM NaOH; pH 2, 10 mM HCl; all solutions contain 20 % MeOH) to give a 400 μM working solution. The solution was filtered through a 0.2 μm membrane filter and analyzed by reversed phase HPLC (see general methods) at intervals (retention time $t_{\text{LLL-}\mathbf{1aaf}}$ = 15.1 min, t_{TPPO} = 17.9 min). The peak areas of LLL-

1aaf from rp-HPLC were measured and normalized against peak areas of TPPO and plotted against incubation time. The data points were fitted to first order kinetics to give the rate constant and half-life $t_{1/2}$ of the decomposition reaction.

General Procedure for the Protease Stability Assay

A stock solution of cyclic peptide LLL-**1aaf** (66 mM) and TPPO (100 mM) in DMSO was prepared and stored at 25 °C. A similar stock solution containing 66 mM linear LLL-**aaf** peptide and 100 mM TPPO was prepared in DMSO as a control. A 0.2 unit / μL stock solution of pronase from Streptomyces griseus was prepared in PBS buffer and further diluted to 0.4 unit / mL with PBS buffer containing 20 % MeOH. 4.5 μL of cyclic peptide or linear peptide stock solution was added to 1.5 mL pronase solution and then filtered through a 0.2 μm membrane filter and the resulting solution was analyzed by rp-HPLC (see general methods) at intervals. The peak areas of LLL-**1aaf** or LLL-**aaf** from rp-HPLC were measured and normalized against peak areas of TPPO and plotted against incubation time. Under the experimental condition, no decomposition of LLL-**1aaf** was observed even after 12 h, while for the control linear peptide LLL-**aaf** the half-life of decomposition was about 1.5 h.

APPENDIX F

WORKS CITED IN THE APPENDIX

(1) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *Journal of computational chemistry* **2005**, *26*, 1781.

(2) Humphrey, W.; Dalke, A.; Schulten, K. *Journal of molecular graphics* **1996**, *14*, 33.

(3) Ko, E.; Raghuraman, A.; Perez, L. M.; Ioerger, T. R.; Burgess, K. *J Am Chem Soc* **2013**, *135*, 167.

(4) Rodriguez, M.; Llinares, M.; Doulut, S.; Heitz, A.; Martinez, J. *Tetrahedron Lett* **1991**, *32*, 923.

(5) Fathia Mosa, C. T., Michel Vaultier, Graham Maw, andrew Whiting *Organic Syntheses* **2008**, *85*, 219.

(6) Marin, J.; Didierjean, C.; Aubry, A.; Casimir, J. R.; Briand, J. P.; Guichard, G. *J Org Chem* **2004**, *69*, 130.

(7) Olivier Chaloin, F. C., Julien Martin, Haixiang Zhang, Gilles Guichard *Organic Syntheses* **2008**, *85*, 147.

(8) Menichincheri, M.; Bargiotti, A.; Berthelsen, J.; Bertrand, J. A.; Bossi, R.; Ciavolella, A.; Cirla, A.; Cristiani, C.; Croci, V.; D'Alessio, R.; Fasolini, M.; Fiorentini, F.; Forte, B.; Isacchi, A.; Martina, K.; Molinari, A.; Montagnoli, A.; Orsini, P.; Orzi, F.; Pesenti, E.; Pezzetta, D.; Pillan, A.; Poggesi, I.; Roletto, F.; Scolaro, A.; Tato, M.; Tibolla, M.; Valsasina, B.; Varasi, M.; Volpi, D.; Santocanale, C.; Vanotti, E. *J Med Chem* **2009**, *52*, 1230.

(9) AbdelMagid, A. F.; Carson, K. G.; Harris, B. D.; Maryanoff, C. A.; Shah, R. D. *J Org Chem* **1996**, *61*, 3849.

(10) Clemens, R. J.; Hyatt, J. A. *J Org Chem* **1985**, *50*, 2431.

(11) Liu, H. J.; Lai, H. K.; Attahpoku, S. K. *Tetrahedron Lett* **1979**, 4121.

(12) Raghuraman, A.; Ko, E.; Burgess, K. *J. Am. Chem. Soc.* **2011**, *133*, 12350.

(13) Raghuraman, A.; Xin, D. Y.; Perez, L. M.; Burgess, K. *J Org Chem* **2013**, *78*, 4823.

(14) Yamasaki, M.; Li, W.; Johnson, D. J.; Huntington, J. A. *Nature* **2008**, *455*, 1255.

(15) Chevallet, M.; Luche, S.; Rabilloud, T. *Nature protocols* **2006**, *1*, 1852.

(16) Mahadeva, R.; Chang, W. S.; Dafforn, T. R.; Oakley, D. J.; Foreman, R. C.; Calvin, J.; Wight, D. G.; Lomas, D. A. *The Journal of clinical investigation* **1999**, *103*, 999.

(17) James, E. L.; Bottomley, S. P. *Archives of biochemistry and biophysics* **1998**, *356*, 296.

(18) Henry, B. L.; Connell, J.; Liang, A.; Krishnasamy, C.; Desai, U. R. *The Journal of biological chemistry* **2009**, *284*, 20897.

(19) Zhao, F. Q.; Craig, R. *Journal of structural biology* **2003**, *141*, 43.

(20) Glenn, M. P.; Kelso, M. J.; Tyndall, J. D. A.; Fairlie, D. P. *Journal of the American Chemical Society* **2003**, *125*, 640.

(21) Beierle, J. M.; Horne, W. S.; van Maarseveen, J. H.; Waser, B.; Reubi, J. C.; Ghadiri, M. R. *Angew Chem Int Edit* **2009**, *48*, 4725.

(22) Steffel, L. R.; Cashman, T. J.; Reutershan, M. H.; Linton, B. R. *Journal of the American Chemical Society* **2007**, *129*, 12956.

(23) Lingard, H.; Han, J. T.; Thompson, A. L.; Leung, I. K. H.; Scott, R. T. W.; Thompson, S.; Hamilton, A. D. *Angew Chem Int Edit* **2014**, *53*, 3650.

(24) Ohnishi, M.; Urry, D. W. *Biochem. Biophys. Res. Commun.* **1969**, *36*, 194.

(25) Cierpicki, T.; Otlewski, J. *J Biomol Nmr* **2001**, *21*, 249.

(26) Duffy, E. M.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2000**, *122*, 2878.